The Effect of Privacy Salience on End-User Behaviour: An Experimental Approach based on the Theory of Planned Behaviour

Thomas Hughes-Roberts

Submitted in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

University of Salford

School of Computing, Science and Engineering

2014

Contents

List of Figures

List of Tables

Acknowledgements

I would like to take this opportunity to thank all those who helped me during this research study. First and foremost, my thanks goes to my partner Vanessa for the unwavering support throughout this process, having two PhD students in the house has not been easy at times but the shared support and experience has been invaluable. Further thanks go to my family for the help throughout my Higher Education journey and life in general, none of this would have been possible otherwise.

Many thanks to the University of Salford and all the staff who have offered advice and guidance when it has been needed; particularly my supervisor, Professor Grahame Cooper for advice, guidance and support throughout my PhD.

My thanks also to all the staff I approached for help with data collection and research involvement; their prompt responses to emails and willingness to participate helped a great deal. In a similar vein, thanks to the students who participated as subjects of the research and gave up their time do so.

Finally, thanks must go to the Graduate Teaching Assistantship program at the University of Salford and to everyone involved from my fellow GTA's to the organising staff. Without the opportunity to be involved with such a program this work would not have been possible.

Abstract

End-User privacy concerns surrounding use of Social Networks present new and complex problems for research. Specifically, a phenomenon known as “the Privacy Paradox” has been observed where end-users stated concerns, attitudes and intended behaviour are not consistent with the actual behaviour within the network. Numerous causes have been proposed as potentially being the root of the problem of this paradoxical phenomenon including a lack of user awareness of privacy issues, a low level skill in using technology or a lack of privacy salience within the social network itself. However, the role of the User Interface (UI) in contributing to, and potentially providing a solution to, poor privacy behaviour is under-explored. A potentially fruitful avenue of enquiry given that behaviour is considered to be a reaction to environmental stimulus and the UI provides the environment within which the user is interacting.

This thesis implements a two phase approach to furthering understanding of privacy behaviour in social networks. First, a survey is implemented exploring the relationship of concepts within the privacy paradox identifying that users stated needs are not being met by their observable behaviour. Secondly, two experiments are implemented in order to explore this behaviour as an interaction with the network; these questions are answered to build a social network profile and can be grouped according to their potential sensitivity. A model of social psychology, the Theory of Planned Behaviour (TPB), is used to develop such experiments in order to examine the cognition behind these interactions. Each of the salient influencers defined by the TPB is used to inform a series of UI treatments and form the basis for experiment groups. An initial experiment explores the method and is used to inform the design of the second, which also introduces a factorial design to explore the relationships between treatments. These experiments show that participants within the treatment groups disclose less information than the control, with statistical significance. Within the first experiment this non-disclosure took place across all questions sensitivities, possibly due to limitations in the experimental method. However, participants in experiment two appear far more selective in their disclosure, choosing not to answer more sensitive questions suggesting that they thought of their privacy while interacting with the system.

Findings within this thesis suggest that the UI plays an important role in influencing end-user behaviour as it can inform the context of the interaction as it happens.

Chapter 1 Introduction – Privacy and the Privacy Paradox

The rise of the Internet and, in particular Social Networking Sites (SNS's), has produced new and complex problems for end-user privacy. Privacy is an already complex social concept which is difficult to define successfully for the individual as the concept itself can be very different from person to person and within varying contexts (Ackerman and Mainwaring 2005). Indeed, the problem has been described as "inherently complex, ill-defined and seemingly insolvable" (Ackerman and Cranor 1999). This makes research into privacy issues difficult as there is no unifying way of thinking of a problem which is different depending on the individual and the context in which the problem exists. The extension of privacy into the new technologies mentioned previously has added to the complexity of the problem for research.

One such example of a privacy related phenomena is the "Privacy Paradox" which describes a disconnect between user's stated privacy concerns and their actual behaviour within online services (Acquisti and Grossklags 2004); where it has been observed within e-commerce sites (Norberg, Horne et al. 2007) and SNSs (Barnes 2006). That is, users are observed using low levels of protection and disclosing large amounts of personal information despite stating they are highly concerned of their privacy when using such systems. However, the paradox itself is under researched where each observation of it examines it in a different way and according to a variety of definitions of it; hence there is a need to unify the research into the paradox and ask the questions: why does it occur within a technological environment and what are factors which influence it? Furthermore, how can the complexity of the privacy concept be explored within the research field in such a way that the field can be unified behind a single definition?

Answers to such questions shall be proposed within this study where a variety of methods are designed and implemented; a survey instrument is designed to provide a full and recent view of the paradox within the context of this study and two experiments designed and implemented with the aim of exploring the role of the User Interface (UI) in understanding, contributing to and solving the privacy paradox.

The phenomenon of this paradox is used to frame this study which is aimed at providing a more complete understanding of the observable aspects of privacy within Social Network Systems answering the above proposed questions. Observable properties of the phenomena are important to this study as they can be quantified and examined for causal relationships to

introduced treatments. Rooting the work in Human Computer Interaction (HCI) allows an exploration of the UI and its role in influencing these observable properties that make up the privacy paradox. Such an exploration could provide an understanding of how UI's can be designed to solve the privacy paradox and encourage more considered privacy behaviour.

1.1 The Role of HCI

HCI is ideally suited to exploring such a phenomenon as it is a cross-disciplined field combining elements of psychology, computer science and User Interface Design (UID). As such, it allows the work to study the behavioural outcomes of the privacy paradox with a view to understanding the effect of the system itself. Therefore, a new understanding of, not only the privacy paradox, but also privacy behaviour in general can be obtained for use in the design of systems which encourage pro-privacy behaviour.

The field is suggested as being uniquely suited in helping the design of systems which satisfy the need to protect sensitive information and can help understand the notions of privacy that individuals have (Iachello and Hong 2007). It is therefore, well placed to shed new light on the privacy paradox through the study of users and their interaction with the system in question (in this case Social Networks). Furthermore, HCI incorporates elements of cognitive theory within it and, through appropriate research methods, models of cognition can be generated which explain computer use (Lyytined 2010) and hence can be used to explain the behaviour seen in the privacy paradox which frames this work providing an understanding of the role of the UI.

Work in the field of visualisation suggests that the User Interface can be embedded within appropriate models of cognition in order to provide assistance with the analysis of data within a computer system (Tory and Moller 2004). Although this is dealing with the visualization of complex data, the idea is easily applied to the arguably equally complex area of privacy conceptualisation. For example, it is suggested that users should be informed of potential risks to their privacy (Fogel and Nehmad 2008) within social networks and HCI provides the means to examine how this should be done and explore the effects of it. Indeed, the Social Network Site Facebook has (after the research work in this thesis was carried out) introduced more salient privacy related information to their system interface (see *figure 1*). Interestingly, the changes to Facebook's UI include some of the privacy salient information proposed within this thesis, which are based on models of cognition. Research has suggested a potential cause of the privacy paradox and poor privacy behaviour in general is a lack of such privacy

salient information in the SNS environment for the users to take into account when interacting with the system (Houghton and Joinson 2010). However, what is privacy salience and how should it be embedded into the UI with a theoretical foundation?

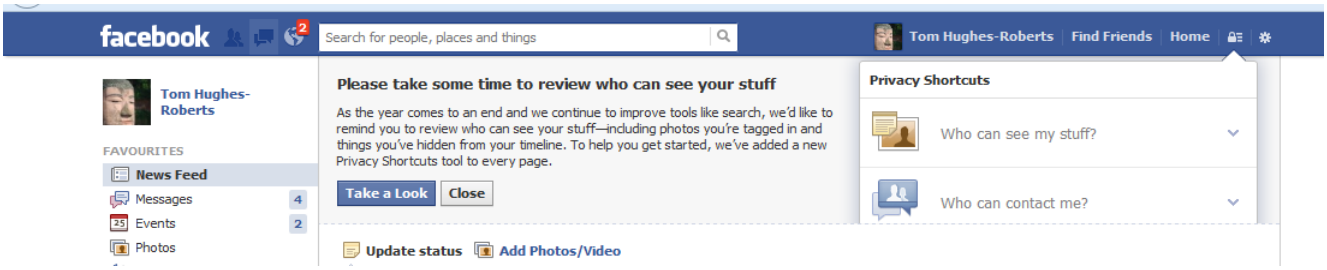


Figure 1 – Facebook UI changes

One final point for the appropriateness of using HCI to tackle privacy problems should be made. HCI allows for the users themselves to be targeted in applying their own personal privacy preferences. Privacy is individually oriented where one person's idea of privacy may not be the same as another person's; furthermore, that idea may shift and change depending on the context at any one time (Palen and Dourish 2003). As such, technical solution to privacy would need to be adaptive enough to cater for the range of individual preferences from potential users. Research has suggested that there may be no technical silver bullet to solving the range of privacy problems which are present in social networks (Rosenblum 2007).

HCI then, has several advantages in tackling the privacy problem. For example, User Interfaces can be adapted to enable "learnability" through their interaction (Johnstone 2003); that is, UIs that promote user learning during interaction. The question is what should be embedded within a SNS UI in order to promote privacy learning and enable the user to implement their individual personal privacy preferences and this where the models of cognition described earlier will be useful.

Users have been described as the "weakest link" where the security of computer systems are concerned (Sasse 2007) and as security and privacy are closely related, it is a reasonable assumption to make that user need to be tackled in a similar way in order for them to safely user the technology in question. Methods of tackling the use of computer systems can be embedded and tested within HCI research; for example, elements of privacy salience that remind users of their privacy safety, information aimed at increasing privacy awareness and manipulation of the way in which sensitive data is presented can all be implemented in the UI. Literature has noted that the role of user cognition is vital in safeguarding user security

suggesting that while practitioners have spent years developing complex systems, psychologists have noted ways in which people consistently misperceive and misunderstand things (Smith 2012). Could it be that social network users misunderstand the scope and complexity of privacy within a digital environment leading to unintended disclosure? Their reported levels of concern would suggest that they are aware of privacy issues but perhaps are not enabled to act accordingly within the network environment.

1.1.2 Are Social Networks a Persuasive Technology?

Within HCI is the emerging field of persuasive technology which suggests that software has the ability to alter the habitual behaviour of users if it is designed to do so (Fogg 1998); and indeed, Facebook has been suggested as being one such technology (Fogg and Iizawa 2008). The research direction for this field of study proposes the use of well-established models from the field of behavioural psychology in line with the research direction proposed for HCI in general mentioned earlier. Hence, the need for models of behaviour within research approaches is demonstrated and this thesis shall provide a review of appropriate models for consideration within HCI and propose an approach to utilising them to examine privacy behaviour. The question then, is what it is within the Facebook UI that persuades users to behave paradoxically (or if, indeed, it is persuasive in the suggested way)? The paradox's existence would suggest that the UI is being persuasive in some way as the behaviour observed is unexpected, that is, it is altered from the way in which the user states that they would normally behave.

Behavioural psychology itself proposes that all behaviour is a reaction to the environmental stimulus within which that behaviour takes place (Breakwell 2006). As such, the role of cognition within the context of the UI as it has been described thus far would appear to be a fruitful area of research for understanding paradoxical privacy behaviour and ascertaining the design features which should be included within a pro-privacy persuasive system.

This work therefore, proposes to examine models of cognition and test the ability to improve privacy behaviour within social networks using HCI designed experiments. Results from such can be used to create general assumptions about privacy behaviour and, hence, a view of the privacy paradox and its causes from a UI perspective.

Proposed within this thesis as a suitable model for informing experiments that aim to understand privacy behaviour is the Theory of Planned Behaviour (Ajzen 1991) which states

that there are three main influencing factors behind behavioural intention and actual behaviour. The model itself is described in greater detail later in this thesis and has several advantages to the research field which make it ideally suited. First, its classification of salient properties into three distinct groups allows for the design of three separate experimental treatments so privacy behaviour can be explored from a variety of angles for a greater degree of richness. Secondly, its experimental foundation as a tool for examining behaviour is well founded providing the study with reliable methods allowing for general and justifiable conclusions to be drawn.

The appropriateness of such will be further outlined in this thesis; first however, the need for the study shall be provided.

1.2.1 Outlining Need

The focus on the "privacy problem" within social networks has been well documented since their introduction and rise in popularity. Figure 2 shows a collection of news articles collected over the first year and a half of this study.

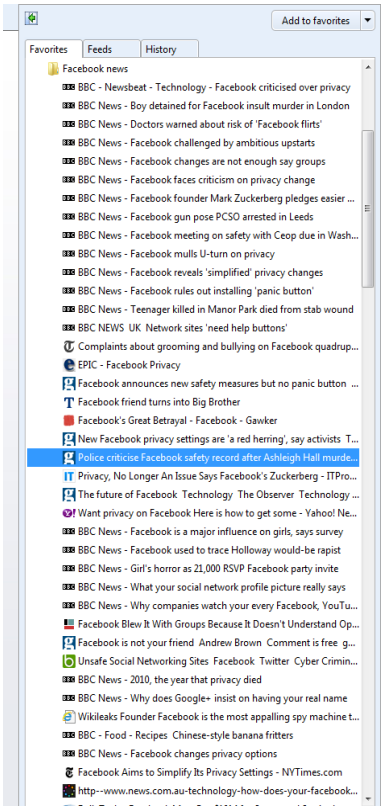


Figure 2 - Example of Media Attention

This is by no means an exhaustive list and reports were only added when they were encountered during daily routine; however, the idea is to give a general overview of the

attention paid to the notion of privacy within social networks. The news articles collected range from concerns over the control afforded by the system over protection of an individual's privacy to the problems that users have experienced in their personal lives due to poor privacy behaviour on their part. Hence, it would appear that there is real world problem with privacy in social networks and that the technology itself is producing an area of concern that has not been seen before.

The paradox then is significant as it is an observable manifestation of this new problem and has significant implications in and of itself. Taken at face value the paradox suggests that an end-user's perceived needs are not being met within the actual network. This could, therefore, result in an individual's personal (and potentially sensitive) information being disclosed to unintended third parties, possibly to the detriment of the user. A range of implications of the privacy paradox are therefore presented that could possibly affect the user.

1.2.2 Implications of the Privacy Paradox

There are numerous implications to the privacy paradox varying in their level of risk to the user. If elements of the privacy paradox are in evidence then it can be concluded that an individual's stated privacy wants are not being met and, therefore, privacy risks within social networks apply all the more. Indeed, some of the reported incidents of perceived poor privacy behaviour can be viewed in the previously illustrated figure detailing news reports. First, users could face embarrassment should their information be disclosed to those who are not the intended recipients of the information (Strater and Lipford 2008). This is potentially a "low impact" implication of poor privacy behaviour yet is still an avoidable unpleasant experience if users control their information flow appropriately. In the same vein of low impact risks a user could suffer annoyance from targeted advertising developed from their disclosed information; an occurrence which has been noted as annoying to consumers in research (Johnson 2010).

Secondly, research has noted that users should be wary of information revelation within social networks due to the increasingly popular practice of checking user profiles of employees by potential and current employers (Gross and Acquisti 2005). Therefore, users who show evidence of the paradox and poor privacy behaviour could face disciplinary action from employers depending on the information they disclose (or potentially not get an employment offer if employers implement social network checks).

Thirdly, the most severe potential implication of the privacy paradox and of poor privacy behaviour in general deals with breaches in the law either toward users or by users themselves. For example, users could be susceptible to identity theft depending on the extent of their personal information disclosure to unintended parties (Donath and Boyd 2004). Furthermore, users could admit to breaking the law through the information they disclose; for example, users who disclose evidence of substance misuse could face legal ramifications if the authorities are unintentionally disclosed to (Morgan, Snelson et al. 2010). The same work suggests that users who engage in this, view the disclosure of such information as a positive thing as they build desired social ties; however, this would only be the case if the appropriate parties are the sole recipients of the information revelation. Further research however, would suggest that this is not the case as users are often unable to manage their information in terms of who can see it (Fang and LeFevre 2010).

However, is the paradox a phenomenon which should be expected? That is, should self-reported concern, attitude and actual behaviour be inline within each other. Acquisti and Grossklags (2004), state that this is a dichotomy that should not be expected; that the sheer magnitude of data from experimental and survey based research shows that users often state levels of privacy concern yet consistently trade-off their privacy for a variety of rewards. Hence, this is still a problem for research to examine given that research as recent as 2010 is still observing instances of the paradox (Stutzman and Kramer-Duffield 2010).

1.3 Summary & Contributions

This chapter has introduced the topic area under discussion and exploration within this thesis and justified the need and approach to studying it.

The questions which therefore arise from the topics introduced thus far can be summarised as the following:

1. What is the privacy paradox and how are its constituent parts related; indeed is it still an issue today?

2. Are social networks persuasive and if so, in what ways are such systems (e.g. Facebook) influencing users (either intentionally or not) with regard to their privacy?

3. Can models of cognition be implemented within HCI research which seek to provide an understanding of the privacy behaviour observed within social network systems?

4. If so, what are the models of cognition and how could they be implemented within a UI to provide users with the information required to make the right privacy choice for their individual needs?

5. Research mentions a lack of privacy salience; if so, what is salient privacy information based on the models of cognition mentioned?

6. What therefore, is the causal relationship between UI elements and privacy related behaviour?

This work therefore, proposes the following contributions to the wider research field;

- A more complete view of the privacy paradox and privacy behaviour in general where social network system use is concerned (note, that although this work deals with the paradox as a framework, the results can be generalised to privacy behaviour in general).

- This is done through the design of a more holistic survey instrument than has currently been used.

- A review of models to be introduced to the field in order to understand that behaviour and draw conclusions from it.

- An application of the elements of behavioural psychology into Experimental User Interfaces aimed at examining the effect of UI features on end-user behaviour.

- A set of results examining the potential causal relationships between the treatments introduced and resulting behaviour.

The rest of this thesis takes the following structure: a literature review examining the concept of privacy and the privacy paradox with the aim of highlighting a way of thinking about privacy within this work and highlighting the need for an examination of the UI. A methodology chapter shall outline the approach taken and the driving philosophy behind it with a following chapter designing the methods to be utilised in this study. These include a survey examining the paradox in a more holistic way than the research field currently offers. Justification for the theories used will follow the results chapter from the survey so as to be informed by the survey results. Finally, the results from the experiments are discussed and conclusions drawn.

Chapter 2 – Literature Review

Chapter 2 – Literature Review – Defining the Problem

2.1 Introduction

The aim of this chapter is to examine a means of thinking about privacy that is conducive to HCI research, to bring together work studying the privacy paradox and to explore the potential causes of poor behaviour that maybe influenced through the UI.

An examination of how privacy has been explored in the field could also highlight the problems of the concept and demonstrate the need for a unifying idea of privacy for use within HCI research. As mentioned by Masiello (2009), privacy is complex and we need a simple way of thinking about it.

2.2 Defining Privacy

The concept of privacy has been described as one "which is in disarray" and one which suffers from a variety of meanings (Solove 2006). Pinning down a single definition for use in research is therefore difficult and the field itself could be skewed from work to work if the concept of privacy driving them differs. Indeed those concepts could be fundamentally in opposition to other definitions of privacy. For example, privacy has been described as a boundary regulation process (Palen and Dourish 2003) and as the right to not be identified (Woo 2006); each of these could produce very different driving ideas for research were one focuses on the privacy being a constant redrawing of lines and the other one deception and anonymity to protect the user.

Each of the above definitions share an ideal of privacy; that it is highly individual and up to the user to implement according to their own needs. Indeed, another definition found within literature finds privacy to be "the ability of the individual to personally control information about oneself" (Smith, Milberg et al. 1996). A similar definition is provided by Westin where privacy is the right of individuals to determine what information about him or herself should be known to others (Westin 2003). These definitions may feel like something of fence sitting ones, deciding that privacy is too complex to define and implement in technology successfully and must instead be left to the individual to implement. There would certainly be some cause to consider such a view with literature describing the concept as a whole as one that is inherently complex, ill-defined and seemingly insolvable (Ackerman and Cranor 1999). The question, then, is how can technology make this sort of provision while maintaining the complexity of privacy?

Within a legal context privacy is normally described as the "right to be left alone" (Levi and Wall 2004); however, this is the antithesis to the very idea of social networks where the goal is to connect and share information with other people. If the concept is then applied to social networks specifically the problem grows and evolves as the context of the situation then becomes a factor in determining the sensitivity of information. For example, what is not considered sensitive at the point of disclosure may become so over time and when taken into account with other pieces of information, something which is possible due to the persistency within social networks. Hence, privacy needs on the web and social networks are time dependant (Lanheinrich 2002) and context dependant (Gandon and Sadech 2004, Masiello 2009).

Considering this point and the individuality of privacy then the conclusion can be drawn that privacy requires a user's constant thought and attention in order to be maintained as they are required to be able to analyse, understand and react to the context they are within at the time. Indeed, it has been suggested that users are required to use their knowledge of privacy to inform an intuitive process of self-disclosure day-to-day (Lederer, Hong et al. 2004). Indeed, a privacy protective mechanism known as P3P, requires users to have a fixed idea of their privacy needs to check against site policies (Ackerman, Cranor et al. 1999); however, further research within the social sciences suggest that users do not have stable, coherent preferences where their privacy is concerned (John, Acquisti et al. 2009). Indeed, is it possible to have stable preferences given the complexity of the concept of privacy as outlined?

There is an inherent difficulty in applying a specific definition into the design of technological solutions as privacy itself is too complex and individualistic to be adequately designed for. Indeed, it has been suggested that there is no silver bullet solution to privacy (Rosenblum 2007) and that instead users must be given the means to implement their own privacy needs. Consider a tutor based system aimed at providing a perfect privacy solution for the individual. Such a system would have to be as complex as the problem of privacy, maintaining an awareness of an individual's context and their needs within a context as it shifts and changes. Clearly such technology is difficult and the users must be relied upon to implement their own desires as they see fit. Why then, within social networks, do users not implement their privacy as they wish?

2.2.2 Privacy in Social Networks

As well as the persistency of data in SNS's adding to the complexity of the privacy context there are other problems which are introduced by the problem domain. Prominent sociologist Erving Goffman described individuals as requiring different social "masks" to cater to the audience in which they are dealing with at any one time (Goffman 1959). So, the individual tailors who they are to suit who they are talking to; an individual may act differently when interacting with their employer than with their friends and this is entirely appropriate. Within social networks, applying such intuitive practices is difficult due to the restrictions of the technology and of the users themselves within that technology where practiced skills and knowledge is required to implement effectively. Given that users do not just add their friends to their social network profiles (Aimeur, Gambs et al. 2009), adding work colleagues, family and even strangers, managing their social spheres present a new problem to the privacy research field. Hence, a requirement of privacy in social networks can be said to be the ability to manage social spheres (Binder, Howes et al. 2009) through identification of relevant spheres and appropriate information revelation to them. This is in some way related to the boundary negotiation process alluded to in other definitions mentioned earlier and is clear example of how this becomes a problem.

Therefore, the element of control becomes ever more important within an SNS in terms of actual information flow (Chen and Williams 2009) as each piece of information is open and available to all depending on the settings applied to it. Users have to be aware of what is disclosed, who it is disclosed to and what could be inferred from it; added complexity is gained when these granules of information could be compiled together to infer something else entirely. However, it has been suggested that granular controls are not utilised sufficiently by users with SNSs (Acquisti and Gross 2006); hence, the environment as it exists is not sufficient in providing users with the control they require. Research suggests that if privacy technology is too complex the features which protects one's privacy are ignored (Grandison and Maximilien 2008). Given the extent of this complexity, how can users be encouraged to make the right decisions?

Research within social networks often states no formal definition of privacy and instead uses the term as a catch-all for privacy risks and issues. In line with the issues discussed thus far, research typically focuses on the information and the users of social networks rather than technological solutions. For example, the "PrivAware" system (Becker and Chen 2009)

analyses potential privacy issues and makes recommendations to the user for improving their privacy. Other research has focussed on the extent of information revelation (Lipford, Besmer et al. 2008) by introducing an "audience" view to allow users to see their profiles as others do. This would suggest that users need to be given extra information and perspective in order to appropriately manage their privacy. Further research explores the use of "friends only" settings as a highly protective form of privacy (Stutzman and Kramer-Duffield 2010). Each of these show varying ways in which the concept of privacy has been considered within social networks; however, without a clear definition of what privacy actually is and an idea of how it manifests itself, makes the design of a method to explore privacy difficult.

2.2.3 Pin-Pointing a Definition

So, the concept of privacy is complex and providing users with a holistic solution that caters for that complexity if difficult. Given that it is also individualistic in nature users must be enabled to make their own decisions amidst that complexity. So where do the problems begin? Solove (2009) provides a taxonomy of privacy:

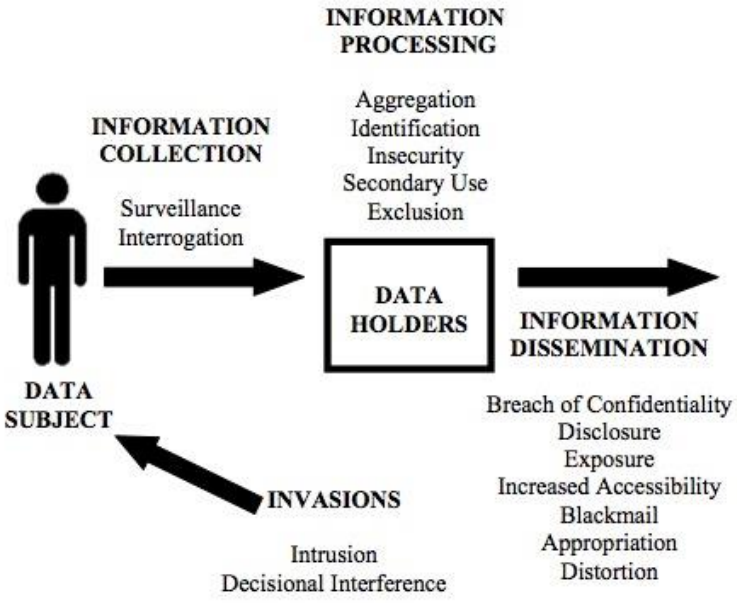


Figure 3 – A taxonomy of privacy (Solove 2009)

This taxonomy of privacy states that all privacy problems begin with the invasion of the data subject or the data subject giving out information about themselves. If this taxonomy is tailored slightly for a social network focus then all potential problems begin with the disclosure of data to the network and, to a lesser extent, with the poor application of privacy

settings which also cause unauthorised parties to gain access to that information. This shifted taxonomy can be seen in the following figure:

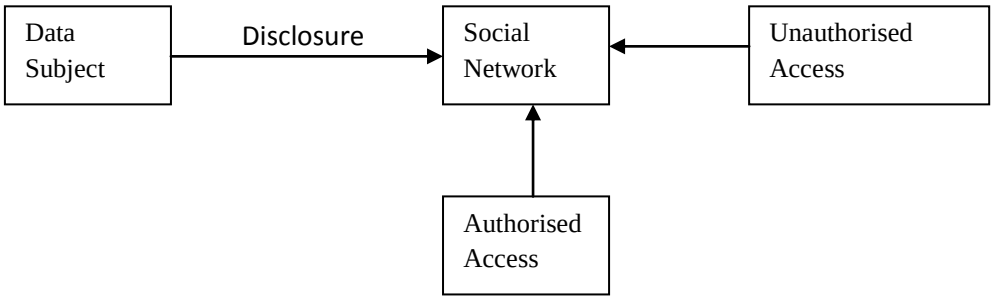


Figure 4 – A taxonomy of privacy

Note, the assertion could be made that access is always authorised to user data in a social network as the user agreed to the site policy and set their own policies for access; however, the paradox states that this is not actually what the user wishes to happen according to their own levels of concern and intention. Hence, the above taxonomy would suggest that all privacy problems stem from the act of disclosure and from poor settings application and the paradox would suggest that users disclose more than they wish and protect as little as they want.

It would appear therefore, that focussing on a solution that addresses what users put on a social network and the settings they apply would be beneficial. Furthermore, both disclosure and settings are discrete aspects of privacy behaviour that users could be reminded of at the point of interaction and what could be the focus of a UI based solution.

2.3.1 The Privacy Paradox

This section aims to bring together literature which deals with the way in which the privacy paradox has been explored and the findings put forward. As mentioned in the introduction the privacy paradox is generally described as a disconnect between a user's desire to protect their privacy and their actual behaviour within the network (Boyd and Ellison 2007).

The paradox has been observed within e-commerce and SNSs as well as online environments in general where each work has used a variety of methods and differing definitions as the parts of the paradox. For example, the aspect of concern has been described as having measures which are too varied across literature (Norberg, Horne et al. 2007) creating a significant challenge to the research field.

Studies within e-commerce examined the potential link between a user's stated level of concern and their recorded behaviour; a survey method was used to measure participant's concern and then compared this to their behaviours within an experimental ecommerce scenario (Jensen, Potts et al. 2005). Concern was given a rating based on the Westin scale of privacy concern, as in similar studies (Kumaraguru and Cranor 2005), where users are given a classification according to their responses of privacy fundamentalist, pragmatist and unconcerned. The assumption being that privacy fundamentalists would demonstrate stricter privacy related behaviours (disclosing less, consult policies etc.). This study would therefore describe a relationship between concern and behaviour where an increased level of concern should correlate to more careful privacy behaviour:

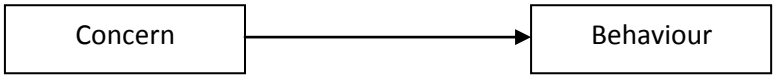


Figure 5 – Concern link

The study found, however, that these two were not negatively correlated. A similar ecommerce study examined the link between reported levels of intention to disclose with actual disclosure using a quantitative survey method (Norberg, Horne et al. 2007). Participants were asked what they would willingly disclose and several weeks later asked related questions in a follow up survey, showing that participants freely disclose information about themselves despite their prior intentions. Hence, the Norberg study assumed a relationship between intention to disclose and the discrete behaviour of disclosure:

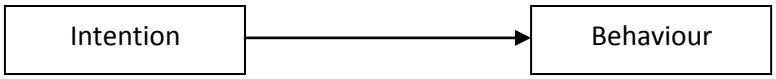


Figure 6 – Intention Link

Again, the study found a the link in figure 6 was not as expected as the two were disconnected as concerned users disclosed information despite their initial, stated intention that they would not.

In each of these studies behaviour is described as the act of disclosing information within the environment. If behaviour is a reaction to environmental stimulus (Breakwell 2006) it is possible that the environment itself is acting upon users to produce the response seen; as it is evident in two studies where two different assumed relations are studied this is a reasonable assumption to state at this stage. For example, a user could go into a system with certain

ideals which are subverted or influenced by the UI (potentially sub-consciously) leading them to act paradoxically.

Moving onto instances of the paradox within social network study; similar to the studies mentioned previously, two pieces of research focussing on SNS environments examined the link between concern and behaviour, where behaviour is information disclosed to the network (Acquisti and Gross 2006, Tuekci 2008). Again, the Westin rating system was used in the Acquisti paper while a more general statement of concern was used by Tuecki. The results from each of these papers are interesting in terms of describing the paradox as, although they seem similar on the surface, each offers slightly varying differences in the findings.

Acquisti and Gross used a survey method to analyse privacy concerns and self-reported behaviour, this was then compared to data mined from the SNS Facebook for each participant in the study. In terms of concern there was no link between the level of concern, membership of the network and the amount of information disclosed. That is, participants with a higher level of concern still joined the network and still disclosed much of their personal information. However, the work found that self-reported information disclosure (this thesis will refer to this as intention from this point forward) matched fairly closely to what information was actually in participant's profiles even if this was not correlated with concern. The paradox was evident in that participants were more visible than they believed themselves to be. The Tuecki paper also found that privacy concern and information disclosure are not related when compared as perhaps might be expected (hence, the paradox).

To focus on the idea of control within a social networking environment for a moment; the Acquisti paper used the profile settings which have been assigned to a participant's network presence. Recall, the paper found that participants were more visible than they thought they were and therefore demonstrated an aspect of the paradox (this did not seem to be examined against their levels of concern). Tuecki expands upon this finding by examining the link between concerns and perceived network visibility using a purely quantitative, survey approach (i.e. no network observations). The work finds that concern is managed through the setting of higher levels of privacy protection. However, as there is no direct behavioural observation of the network itself it is unknown if these reported settings are evident in actuality. The paradox would suggest that they would not be and, indeed, the Acquisti study proposed such as part of its findings.

Before continuing a brief review and conceptualisation is required; first, the Acquisti paper explored the link between intention and disclosure, finding that the two matched up well:

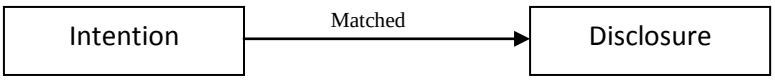


Figure 7 – Acquisti Study's Assumed Relationship

Notice that this is a similar concept map as illustrated through the Norberg (2007) paper mentioned previously but the two found different results. Norberg's data pointed to the two being inconsistently linked, unlike Acquisti's, showing the variance in research within this field.

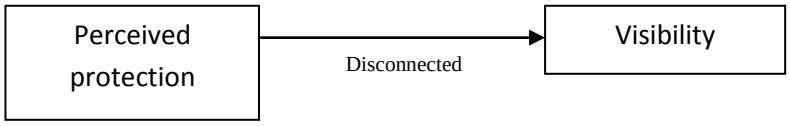


Figure 8 – Paradoxical Variation

The study also found that the perception of visibility did not match how visible participants actually were in terms of the protection settings applied; hence their intended privacy settings did not match the actual settings applied.

Tuecki, meanwhile, explored the link between concern and both disclosure behaviour and visibility finding that participants *reported* that they protect themselves more if they are concerned yet this did not play any bearing on the amount of information disclosed:

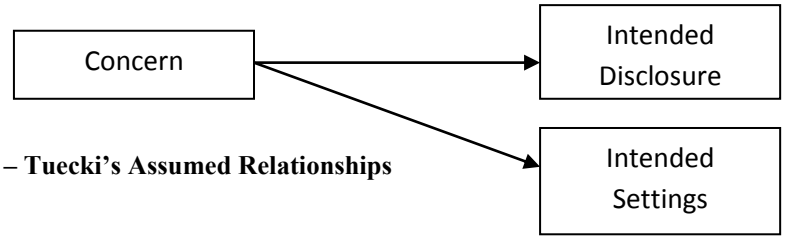


Figure 9 – Tuecki's Assumed Relationships

Again, the above conceptualisation represents an assumed relationship; higher concern should lead to a higher intention to non-disclose and protect more. As this was dealing with self-reported behaviour each of the two elements studied can also be described as *intention* to disclose and *intention* to be visible.

Each of these studies would suggest that the application of privacy settings and disclosure behaviour are two separate entities holding a different relation to concern in terms of the

findings offered. Perhaps this is to be expected as an individual may not be concerned about their privacy if they only believe their close friends to be the recipients of their information. Indeed, there is literature that suggests that disclosure and control are not negatively correlated (Christofides, Muise et al. 2009); the results of this came from a pure survey approach, again with no direct network observations. This link between control and disclosure is further explored through experiments in an online survey where increased control over data was shown to increase levels of disclosure which would suggest that there is a relationship between the two (Brandimarte, Acquisti et al. 2012). Here less control decreased disclosure and increased concern despite risks decreasing. Already here, it is beginning to become clear that research in the area is mixed in terms of findings offered; although all studies demonstrated paradoxical information they all came from within different places of the elements studied and occasionally contradicted other studies. For example, ecommerce research (Norberg, Horne et al. 2007) found that participants disclosed more than they thought they did, while Acquisti (2006) stated that participants knew how much they disclosed but the amount was not related to concern.

A more recent work suggests that users do implement their privacy settings to manage their privacy expectations, but struggle to manage their friends circle within the network itself (Johnson, Egelman et al. 2012). This is at odds with a separate work that found that despite users saying they understood and implemented comprehensive settings, they actually did not (Debatin, Lovejoy et al. 2009). Clearly, then there is a need to further examine this phenomenon and understand the actual behaviour in question. If the application of settings themselves is inadequate in addressing privacy needs then perhaps more careful disclosure habits should be considered. Such shall be the focus of this thesis.

To bring together the literature studied here, figure 10 offers a summative model of the expected relationships where privacy is concerned; concern, attitude and behaviour were suggested as potential relationships by Jensen, Norberg and Acquisti, while, behaviour can manifest itself in terms of the granular, discrete observations of disclosure and settings applications within the network.

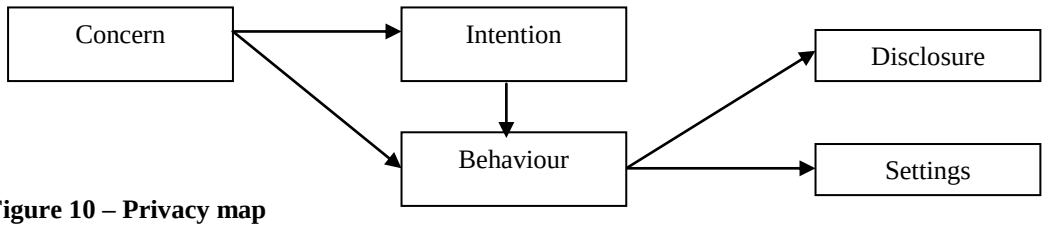


Figure 10 – Privacy map

Evidence for the paradox was found where these causal links where not as expected; note that all studies examined privacy in terms of the two behaviours suggested by the privacy definition section. Not one study examined the paradox in the entirety illustrated here or used a variety of measures to do so. Furthermore, empirical evidence for privacy behaviour is also severely limited, as behaviour, in relation to privacy, is granular in nature. That is, SNS privacy deals with individual pieces of sensitive information, with each behaviour being an isolated interaction with a data granule (Krishnamurthy and Wills 2008). Granular control is how behaviour is exercised within a social network (Stutzman 2006), so participant perceptions and actions at this level are relevant to a more holistic use of the system.

Figure 10 matches the aspects of the taxonomy in figure 4 in that the observable manifestations of the paradox are disclosure and settings. However, it is unclear whether the paradox occurs between concern and behaviour or intention and behaviour and as such further work is required (as either or both can be influenced by the UI).

2.3.2 The Causes of Poor Behaviour

The role of user awareness has been questioned as an influencing factor of the privacy paradox. For example, one such paper examining the "posting paradox" states that users disclose information in spite of awareness that inappropriate parties may be viewing their data (Miller, Salmona et al. 2011) and proposes that disclosure can be reduced via increasing privacy concern. However, the paradox shows users tend not to act according to their personal level of concern. So the question is raised: how should concern be raised and what exactly is the kind of concern that needs to be treated in order to positively affect behaviour?

Awareness is also proposed as a solution, specifically to the paradox, from a separate proposal paper (Pötzsch 2009) which offers the view that users forget their privacy concerns and knowledge during system use and reminding them is a way of closing the disconnect. Interestingly, the paper recommends solutions taking into account the cognition and behavioural aspects of privacy to support the intentions of users. However, the paper suggests that such a tool for raising awareness would need to have knowledge of shifting contexts, something which this thesis proposes adds too much complexity to an already complex problem. This work agrees with the view that reminding users at the point of interaction could encourage consistent privacy behaviour and goes further in suggesting that awareness is too

vague a notion to simply apply successfully; awareness of what precisely, what of instances where intention matches one aspect of behaviour but not another, etc.?

To demonstrate further; users have been described as being unaware of privacy issues and inexperienced regarding the concept in general (John, Acquisti et al. 2009, Kolter and Pernul 2009). They could, also, be unaware of the "openness of the Internet in terms of who can view their profiles" (Barnes 2006) or due to a lack of comprehension, awareness or concern for privacy within the context of an open networked system (Stutzman and Kramer-Duffield 2010). Clearly then, the cause from a user perspective is varied even within research and lacks any empirical basis for formally exploring the problem satisfactorily. Would an increased awareness improve behaviour? Would an increase in skill level or a guide to privacy settings improve privacy protection? Understanding the cognition behind behaviour is vital in answering such questions. For example, hyperbolic discounting suggests that users favour short term gain over long term risk and this has been suggested as being a cause of poor privacy behaviour (Acquisti and Grossklags 2004). Although these relate to the user's context (how much they know, how much they are swayed by cognition), they are also coupled to the system and the UI as this has the potential to increase awareness through design. Would user behaviour be influenced if they are made aware of the extent of long term risk explicitly during interaction?

Moving away from the user's individual awareness brings the review to external influences. Heightened feelings of concern can be elevated by sensationalist news reports and increased media attention (Norberg, Horne et al. 2007); hence, the reported levels of high concern in survey data could be caused by users knowing what response they *should* have. Demonstrated here is the problem of studying a phenomena by separating it into components (behaviour and concern), each with their own context which could be influencing responses and action independently. Behaviour, for example, could be influenced by peer-pressure and herding behaviour (Strater and Richter 2007) which increases the amount of information disclosed within a social network. This is also supported through the theory of social capital which suggests that users in social networks would disclose particular pieces of data in order to build social ties with particular peers (Valenzuela 2009).

Finally, and most relevant to this study, factors based in the system have been proposed as a potential cause of the privacy paradox. Social networks themselves are designed to be open (Fogg and Iizawa 2008, Livingstone 2008), indeed their business model depends on it;

however, how they are designed to be such is unexplored within literature with empirical backing. Some work suggests a lack of privacy salience within the environment so privacy is not a part of the privacy making process (Houghton and Joinson 2010). What salience is and what form it should take is unclear and shall be explored further in this work. Furthermore, it has been suggested that the aspect of granular control is too complex for users to appropriately implement (Fang and LeFevre 2010) and that users require aid in understanding the complexity of the network environment they are using. This could be tied to the idea of signal detection theory which describes the extent to which a "noisy" environment prevents an individual from making "correct" decisions (Heeger 1998). This could explain the potentially related phenomena of the control paradox (Bandimarte, Acquisti et al. 2012) where increased control leads to increased disclosure of personal information. Hence, it would seem that the technology itself plays a significant role in influencing user behaviour.

Cause	Description	Literature
A lack of privacy concern	Users are aware that unintended parties may view their data and are not concerned about the risk	Miller *et al*, 2011
A lack of privacy awareness at the point of interaction	Users "forget" their privacy concern when disclosing information and do not use it in the decision making process	Potzsch, 2009
A lack of awareness of privacy issues	Users are not educated about the risks of privacy or of the concept when placed in a technical setting	John, Acquisti *et al*, 2009 Kolter and Pernel, 2009
Hyperbolic Discounting	Users forego long term risk in favour of short term gain.	Acquisti and Grossklags, 2004
Peer pressure and herding behaviour	Users do as their friends do in order to fit in	Strater and Richter, 2007
Media Attention	Increase media attention may artificially inflate concern giving the results seen in privacy surveys	Norberg, Horn *et al*, 2007
System persuasion	Social Networks are designed to be open and encourage poor privacy behaviour	Fogg and Iizawa, 2008 Livingstone, 2008
A lack of privacy salience	The system does not include privacy information to be used in the decision making process	Houghton and Joinson, 2010
System complexity	The granular nature of privacy control is too confusing for users to successfully adopt and confuses the system interface	Fang and Lefevre, 2010 Bandimarte, Acquisti *et al*, 2012

Table 1 - Assumed Causes of the Privacy Paradox and Poor Behaviour

This brief review of the causes of the paradox (summarised in table 1) demonstrates the complexity of the problem and the lack of defining research within the field. Also missing is the role the environment itself plays, although this is hinted at by the references in the

previous paragraph there is no formal declaration that the UI could play a vital role both causing and finding a solution to the paradox.

2.3.3 Further Impact of the UI

Given that behaviour can be considered a reaction to environmental stimulus (Breakwell 2006) and UI provides that environment then it is reasonable to assume it can play a role in causing and solving the paradox. Indeed, it has been suggested that Facebook is a persuasive technology, designed to alter habitual user behaviour (Fogg and Iizawa 2008). How is it designed such and how can it be designed to achieve the opposite?

It maybe that defining privacy salient UI features that influence users could be a viable and potentially simple solution. It has been suggested that the UI must be designed with societal issues in mind in order to encourage user engagement with those issues (Lederer, Hong et al. 2004, Hochheiser and Lazar 2007). How should these societal issues inform design and what should they look like?

Users may be driven by some pre-conceived goal when interacting with computer systems and it has been suggested that user will complete numerous sub-goals in pursuit of a single goal possible to their detriment (Jacko and Sears 2003). Indeed, privacy has been suggested as being a secondary goal problem (Bonneau and Anderson et al. 2009) so users will not consider it during interaction. How can it be made a primary goal?

There appears to be a need for greater focus on the UI design where privacy is concerned (Masiello 2009). It is well placed to provide solutions in its role as an environmental stimulus; particularly as it has been suggested that altering intention through intervention strategies has a weak influence over actual behaviour (Webb and Sheeran 2006). Furthermore, it has been found that behaviour is more likely to be altered as it is happening rather than through increasing general awareness (Camp, McGrath et al. 2006).

There is a need however, for a greater notion of cognition in order to address the increasingly complex problems facing HCI today (Lyytined 2010) such as the privacy paradox. Such an understanding can inform the design of UIs that address the causes of poor privacy behaviour illustrated in table 1. This research, then, proposes to provide a greater understanding of the paradox through the following contributions:

1. An examination of the figure 10 as a complete view of the privacy paradox.

2. An examination of appropriate models of cognition that explain privacy behaviour found in social network users.

3. A definition or privacy salience that addresses the causes listed in table 1 and that can influence the discrete privacy behaviours outlined previously.

4. An interpretation of that salience into User Interface features.

5. An exploration of the potential effect of that privacy salience in controlled experiments.

2.3.4 Other Related Work

Finally, a brief look at some related work and a statement of where this work fits within the HCI and privacy research field. Related work is considered as pieces of research where the aim has been to examine the effect of the UI changes on related privacy behaviour.

The work of Ackerman and Cranor deals with the creation of privacy tutors to aid in the privacy awareness of end-users (Ackerman and Cranor 1999, Cranor and Reidenberg 2002, Cranor, Guduru et al. 2006). The results obtained from this work and the conclusions drawn will provide empirical evidence for the effect these tutors could have and will inform the design of such tutors, providing a way forward for how they can be embedded within the UI of systems. For example, Cranor *et al*, (2006) proposed the "Privacy Bird" which aims to match user privacy preferences with web site privacy policy. This work aims to examine how to design such a tutor which does not require users to have pre-set preferences but instead aims to remind users that they should think about their privacy when deciding to disclose information (and examine what effect this will have). Hence, there is a need for a pre-cursor to such tutor systems which inform and allow users to develop and identify what their privacy preferences are.

So with relation to work examining the role of environmental changes on end-user behaviour the following are taken as examples to show that there is the potential for change. First, the control paradox resulted from changes to the UI, were extra elements of control were added resulting in an increase in levels of disclosure (Brandimarte, Acquisti et al. 2012). This was observed through altering how questions were asked in online surveys and altering the perceived control over the resulting information requested; participants could either publish themselves or allow an unnamed researcher to control the resulting information. The work

found that more control over the publication of information led to a decrease in privacy concerns and an increase in their willingness to disclose information demonstrating that changes in the network being interacted with can result in changes in behaviour and perception. While the Brandimarte study is similar to the one proposed within this thesis (a faux social network used), the reasons for the observed effects are not covered with empirical validity. The approach proposed within this work aims to examine the role of the UI with a foundation in cognitive models, thus explaining the reason certain UI elements have an effect on users of social networks.

A further example of some related works is that of Hui *et al* (2006) and LaRose *et al* (2007); each of which examined the effect of the presence of privacy seals and privacy statements on privacy behaviour. The first of these studies placed these elements on a company's website thus using a field experiment to examine their effects, finding that a privacy statement induced increased levels of disclosure, as did a monetary incentive and finally, that an information request had a negative effect on disclosure (Hui, Teo et al. 2006). The second of these studies examined the effects of privacy warnings on user's privacy perception and behaviour. These warnings were "clear, conspicuous and concise presentations of the benefits and risks associated with database information practices" (LaRose and Rifon 2007). The work found that warnings increased perceptions of risk and decreased disclosure unless a privacy seal was also present and therefore recommends that development of privacy warnings as an appropriate method of tackling poor privacy behaviour.

Work within personalization has found that users can be grouped into "disclosure groups" where some users are more likely to disclose certain kinds of information but not others (Knijnenburg, Kobsa et al. 2013). The work further suggests that a recommender system could remind users of their particular preferences based on which group they fit into. Further work explored the use of justification messages, describing the reasons for and the benefits of disclosure within a recommender system finding that they did not increase disclosure but did decrease satisfaction and trust in the service (Knijnenburg, Kobsa et al. 2013). This would suggest that additions to the UI must be carefully considered and themselves not off putting to users. Indeed, statements of privacy are usually not read and privacy seals not well understood according to surveys and so, there is no silver bullet for enhancing the privacy friendliness of a system (Kobsa 2007) so small enhancements are necessary.

One final example of related work examined the effect of additional options when information is requested by a service (Joinson, Paine et al. 2007). The work added a "prefer not to say" option when asking questions among the appropriate responses finding that this lessened the amount of disclosure where it was present. This was noted as a form of privacy salience as it is embedding some privacy related information into the environment. Furthermore, a chance to "blur" information was also offered (e.g. giving an age range rather than a specific age) which also reduced levels of specific disclosure, particularly among males. Again, this is an example of how changes to the UI can invoke a behavioural response from end-users providing justification for this study. The cognitive reasons will be explored adding to the level of understanding which is currently available regarding the privacy paradox and privacy behaviour in general. This is an important point to note, although the privacy paradox is framing this study, the results exploring it will also give a broader understanding of behaviour in general.

Chapter 3 – Methodology

Chapter 3 – Methodology

This chapter is concerned with detailing the research philosophy driving this work in terms of answering the questions posed in the literature review chapter and filling the gaps which exist within the field. These questions are aimed at further understanding the privacy paradox through the study of its constituent parts and the relationship they hold with each other with the aim of building a view of the entirety of the phenomenon. This would inform a study of behaviour and its relation to the User Interface within which the behaviour is performed and reacted to; again, as the role of the UI has not been fully explored with a theoretical foundation.

The *type* of questions this research asks are vital in identifying an appropriate research philosophy so the aims of the questions can be answered fully. This chapter takes this into account, as well as the researchers' own personal philosophy and that of the field the research is taking place within.

3.1 Research Philosophy

The research philosophy dictates the way in which the research questions are explored and answered; providing justification for the methods implemented, data gathered and the way in which that data is then analysed. Philosophies can be broadly split into two main camps: Interpretivist and Positivist with each underpinning different approaches to exploring research questions. Positivists hold the belief that the world can be divided into quantifiable observable phenomena which can be measured objectively; understanding of these simple, observable laws can explain the complexity which they often produce. Hence, the bigger picture can be explored by reducing it to the manageable and measurable factors that make it up.

The second philosophy, Interpretivism, holds that many interpretations of reality are possible and that reality can only be understood by acknowledging, subjectively, one's own interpretation of the phenomena being studied. Furthermore, that interpretation can be constructed into scientific knowledge of the problem being researched. This tends to focus on aspects of phenomena which are not measureable but instead understanding must be constructed from subjective interpretations of complex and, usually, qualitative data (Oates 2006).

3.2 Research Questions and Philosophy

Taking an initial look at the research questions and the definitions of research philosophies shows that a positivist approach to this study would be most aligned. In the first instance, understanding the privacy paradox, the relationships between the identified components need to be identified and explored. As such, a positivist approach allows for the individual components to be classified as appropriately measurable, allowing for statistical examination to identify causal relationships. For example, measurements of concern can easily be compared to measurements of intention and the relationship quantified in terms of statistical relevance (i.e. if there is a relationship or not). The difficulty lies in identifying an appropriate measure of subjective phenomena which the following chapter seeks to resolve. Furthermore, empirical evidence for the presence of the paradox in general can be provided through such an approach and the disconnect identified with scientific certainty.

Looking at the second question shows that a positivist approach is also necessary in exploring the role of the User Interface in producing privacy behaviour. The causal relationship between the UI (and specified UI elements) and any privacy related behaviour can be measured, demonstrated and test against hypotheses; again, this would show that there is empirical evidence for the role of the UI in influencing the privacy paradox. The nature of the study shows this is an easy fit as this research is, ultimately, concerned with clearly defined behaviours which are both measurable and observable.

So, the main aim of the research is to examine *what* UI elements produce *what* behaviours more frequently and a positivist approach would allow such an aim to be achieved with comparative measurements of defined behaviours.

Indeed, the hallmark of HCI research is rooted in behavioural psychology (Lazar, Feng et al. 2010) and typically follows the positivist approaches taken within that field (Haslam and McGarty 1998). This is ideal for this study in particular given that it is, fundamentally, an observation of behaviour within a social networking system.

3.3 The Role of Reductionism in Research

This brings the chapter to an important point regarding the philosophy of the work; the role of reductionism. The previous chapter made clear the complexity of privacy and the concept itself has been described as difficult to research within (Ackerman and Cranor 1999). Examining privacy in the way proposed in this thesis gives the advantage of reducing the

problem to a manageable snapshot where privacy is represented as the behaviour that it manifests as. Hence, assumptions about the wider concept can be generated from the empirically true data gathered here. This research then is concerned primarily with privacy *behaviour* and not with the complex concept itself; furthermore, the work is dealing with behaviour within the bounds of the *privacy paradox*. However, assumptions can be inferred about behaviour in general and about privacy in general from the results presented. Indeed, such a view of studying privacy has been noted as typical within the research field of HCI (Dourish and Anderson 2006).

3.4 Initial Limitations

A purely positivist approach to studying a research problem has well documented limitations which could be placed upon the work. Namely, that the why of the positivist results are unexplored; for example, while the main aim of the work is to identify *what* behaviour is related to *what* UI elements and to identify *what* the relationships are within the privacy paradox, *why* those relationships exist and *how* the user feels about them is not covered within the philosophy driving the work. So, an appropriately designed UI element could affect behaviour (where both affect and behaviour are measureable); however, why there is an effect is not covered within a positivist approach.

Therefore, where possible, a mixed method approach should be implemented; that is, while the main driving philosophy behind the work will be positivist in nature (as the main question demands it) some interpretivist approaches shall also be utilised. Such an approach has been highlighted as necessary (Russo 2000) in the past where descriptions of practice must be coupled with interpretations of it.

3.5 Initial Summary

This initial review of research philosophies in relation to the specific questions of this work has shown that a positivist approach is favoured in exploring the topic area. This is summarised in the following points;

1. The work is primarily dealing with defining and providing relationships within the privacy paradox as put forward by literature.

2. This also includes examining the effect of UI elements on behaviour.

3. Each of these are measurable and, furthermore, require measurements for a definitive answer.

4. Hence, a positivist approach is favoured in driving the main arguments of the work.

5. However, due to limitations of a "pure" approach, where possible mixed methods should be implemented.

These five points summarise the main arguments of this section. The following sections describe the methods generated from taking such an approach. Points initially raised here regarding benefits and limitations of the approach used shall be expanded upon and address within the appropriate sections.

3.6 Overview of Methods

This section shall expand on the methods available and the methods to be selected for use in exploring the problem areas illustrated in this thesis. The following table illustrates the typical approaches found within positivism (Galliers 1991):

Table 2 - Positivism

Positivism
Experiments
Field Experiments
Surveys
Case Studies
Theorem Proofs
Forecasting
Simulations

The first area to be explored is the conceptual model of the privacy paradox which is important in informing and underpinning the work that follows it; that is, examining the role of the UI in the phenomena.

The goal of this conceptual model and the research surrounding it is to examine the phenomena in greater detail and more holistically than other works have done so thus far. This examination is required to test the proposed relationships proposed by literature and

identify any causal relationships between them; e.g. does concern match intention and how is this related to specific and observable behaviours.

A positivist approach offers an ideal method in examining these relationships through a survey based method (as can be seen in the above table) which identifies patterns within samples taken in a systematic and standardised way (Oates 2006); hence, relationships can be quantified and evidence for the paradox empirically justified (Gable 1994). Surveys within HCI are typical and can most effectively capture an overview of system usage and how users are interacting with it (Lazar, Feng et al. 2010). In terms of this study, this system use can easily be compared to user perceptions and intentions; thus examining the relationships proposed to exist within the privacy paradox. Such a survey instrument shall be developed based on existing research (Govani and Pashley 2005, Acquisti and Gross 2006, Stutzman and Kramer-Duffield 2010, Brandimarte, Acquisti et al. 2012) with the aim of examining the privacy paradox in greater detail, bringing together literature within the field and informing further study.

Surveys are often used within behavioural research also where they can be used to gain a snapshot of what participants are thinking at one point in time. Furthermore, surveys are useful in measuring the relationships between variables in order to examine attitudes and behaviours over time (Cozby and Bates 2012). Each of these points shows how well the method in question pairs up with the research problem at hand. A snapshot of the problem may seem limited, however, the complexity of privacy has been noted and by producing an observable snapshot of the problem it is put into a form which can be studied in such a way that allows for falsifiable assumptions to be made. However, as surveys deal with self-reported behaviour and not actual, this work seeks to examine its efficacy in exploring the complex concept of privacy and by extension the privacy paradox.

In keeping with the fields which drive this work (HCI and behavioural psychology) appropriate experiments examining User Interface elements forms the second and major part of this research; namely, the role the UI plays in influencing and contributing to the privacy paradox. Experiments are essential within HCI to examine user reaction to the technology being studied (Lazar, Feng et al. 2010) which is clearly ideally related to this research. Similarly, controlled experiments are a hallmark of research within behavioural psychology as causal relationships between treatments and observed behaviours can be identified and measured (Breakwell 2006). Again, this is in keeping with a positivist methodology, as

illustrated by the above table, where the research is focusing on the examination of causal relationships between certain UI elements and the privacy behaviours. In this case it is behaviours of personal information disclosure and the application of certain privacy settings onto that information.

Such an approach allows for statistically reliable results for the sample taken and, if validity is assured, are immediately relatable to a variety of contexts (Haslam and McGarty 1998); hence, the complexity of privacy can be studied through internally and externally valid experiments offering a "true" snapshot of a particular aspect of privacy. In reference to behavioural psychology, behaviour specifically is ideally explored through pre and post experiments comparing observable behaviours across appropriately conceived sample groups (Somekh and Lewin 2009). Such experiments are required to be designed such that observed differences are a result of any treatment applied to the groups under study (Acquisti and Grossklags 2004). Therefore, the treatments, which are the salient privacy features designed by the research, need to be clearly defined with justifiable choices made regarding their makeup and composition.

However, it was mentioned that a purely experimental approach driven by positivism has its limitations; namely a lack of exploration of the *why* observed behaviours occurred within the experiments according to the users perceptions. Such an approach could highlight dissonances between behaviour and understanding. For example, cognitive dissonance, where people explain conflicting views they may hold (Lee, Jung et al. 2011) can only be identified if perceptions are identified which differ from the behaviour and examining how those perceptions are explained by the participant. Therefore, where appropriate qualitative interviews shall also be conducted after the experiments in order to gather a richer set of data aiming to user thoughts towards to the experiments they participated in. Furthermore, during the experiments, observations shall take place examining how users interact with the systems put in front of them. The addition of these extra methods to the research makes the approach taken a semi mixed-method one; such an approach is recommended as appropriate within system work to allow for data triangulation, i.e. for one method to corroborate or add to the findings of another (Oates 2006).

3.7 Methodology Summary

This section has outlined the broad approach being taken by the research to answer the questions posed in the literature review. The work is to be driven by a mainly positivist approach which allows for the measurement of perceptions and behaviours to allow them to be easily and objectively compared to each other. This ideology is carried into the second portion of work where behaviour is objectively compared to a set of pre-defined UI elements acting as group treatments in a set of experiments. Observations and interviews shall also be added to this section to add a greater degree of data richness to the work and allowing for more reliable conclusions to be drawn from the resulting data sets. Hence, the work shall be gathering a mix of quantitative (from positivist approaches) and qualitative (from the interpretivist interviews) approaches. The details and design of these approaches shall be discussed later in the thesis in appropriate chapters.

At this stage, some questions remain open; for example, salience is again mentioned in this section. However, what it is and what it looks like remains unclear. This thesis shall need to find a justifiable way of identifying and dealing with this concept to aid the design of the experiment. Furthermore, the appropriate design of privacy salient UI features should aid in the drawing of sound conclusions from the data should they be embedded into experiments with validity. They are therefore, vitally important to this work and will be the theory upon which the experiments in particular will depend.

The following chapter begins exploring this point by outlining the survey and designing an appropriate instrument for use in examining the privacy paradox and in informing the design of experiments. Following the survey a review of the field of behavioural psychology shall be conducted in order to create a framework for the assumed causes of the privacy paradox within a well tested theory. From this, concepts of salience can be dawn, thus informing the design of the experiments and analysing the data from the surveys in a novel way. Furthermore, an appropriate behavioural theory should allow for more justifiable conclusions to be dawn as any observable behaviours can be explained using the composition of the well-tested theory.

Chapter 4 – Survey Design

Chapter 4 – Survey Approach Design

This chapter deals with the design of the survey instrument aimed at exploring the privacy paradox in greater detail than currently exists within literature and to ascertain that the paradox is still in evidence within the sample to be chosen within this study. To review then, the literature review section dealing with the paradox identified a conceptual model of the phenomenon based on the elements studied within the research field. These elements included concern, intention and behaviour (disclosure of information and the application of settings) where a disconnect can exist between any. Using this model as a basis for exploration the survey instrument has the following aims;

1. To identify where precisely the phenomenon of the paradox exists.
2. To provide statistical evidence for this phenomenon
3. To examine behaviour granularly as it exists within social networks and as it is enacted upon by end-users.

So, this survey instrument is aimed at gaining a more complete picture of the privacy paradox using figure 10 as the basis for exploration. In keeping with, and advised by, a positivist philosophy the survey shall quantify the relationships between elements so causality can be examined. Specifically, figure 10 suggests that concern should hold a relationship with attitude and behaviour (observable as disclosure and privacy settings applied). As such, measures are required for concern, attitude and perceived behaviour and a follow-up measure of actual behaviour.

This chapter then shall deal with the design of the instrument itself, examining each element requiring a measure and discussing an appropriate selection of questions for use in the final instrument. Furthermore, the approach taken to the research shall be discussed with a statement of method made in terms of sample selection and analysis techniques to be implemented. Finally, limitations to the method and approach taken shall be mentioned with appropriate acknowledgements made of the issues to be aware of in this regard. The full instrument developed as a result of this chapter can be found in appendix 2.

4.1 Survey Design

So, the purpose of the survey is to measure concern/attitude and intention (observed behaviour is also examined within this approach but not by the survey; this is discussed later

in the chapter). Hence, measurements are required for concern where the participant's view of privacy is measured, for attitude where their views of social networks are defined and of their intention where reported behaviours are examined in terms of privacy settings and granular data control.

4.1.1 Assessing Participants' Privacy Concern

Typically within research assessing privacy concern the Westin measurements (Westin 1991) are implemented (Ackerman, Cranor et al. 1999, Stutzman 2006). These measures have evolved over the years into a variety of privacy related questions assessing concern differently based on the problem area at hand (Kumaraguru and Cranor 2005). For example, in an extensive study of the concepts of concern and attitudes while online (Cranor 2000), the Westin privacy rating was used to measure the concern of participants in the study. This measurement is a collection of three questions, the answers to which group participants into three clusters of concern (unconcerned, pragmatic and fundamentalist). A review of studies where this instrument was used found that samples of participants were typically broken down as 18%, 57% and 25% respectively (Kumaraguru and Cranor 2005). Similar approaches have been utilised in several papers where part of the aim of the survey instrument is to measure concern (Consolvo, Smith et al. 2005, Jensen, Potts et al. 2005). Consolvo's work examined the use of the measure as an indicator of location disclosure behaviour, while, Jensen's (as mentioned previously) used it to explore paradoxical behaviour. Hence, its applicability within this research field is demonstrated as it has a history of use. The following are the questions as conceived by Westin:

1. Consumers have lost all control over how personal information is collected and used by companies.

2. Most businesses handle the personal information they collect about consumers in a proper and confidential way.

3. Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.

For each question, participants answered by indicating their agreement to the statement from strongly disagreeing to strongly agreeing in a 5-point Likert scale. Should participants agree in any way with the first question and disagree any way to the second and third they would

receive a rating of privacy fundamentalist. Unconcerned participants would answer the opposite and pragmatists are the remaining participants.

As this is considered a typical method for measuring concern this research shall also implement a version of the Westin question tailored to online study (Westin and Interactive 1999). This will maintain consistency with the larger research field and also is required as this aim of this survey is to bring together research dealing with privacy and behaviour.

The following set of questions is indicative of the Westin set utilised within this survey instrument:

1. Users have lost all control over how their personal information is collected and used by social networking sites.

2. Social Networking sites handle personal information they collect in a proper and confidential way.

3. Existing laws and site policies/practices provide a reasonable level of protection for user privacy today.

Again, the same clusters apply where fundamentalists are thought to be strictly private and exhibit this through their behaviour (e.g. low levels of self-disclosure), pragmatists are more flexible in their privacy outlook (e.g. behaviour fits the context) and unconcerned individuals have a lax outlook towards privacy (e.g. protecting little and disclosing much).

However, the literature review made a point of the complexity of privacy and hence, the complexity of the concept of privacy concern. As such, this survey instrument should use multiple measures of concern for a deeper exploration of the privacy problem as perceived by the end-user. Therefore, a second measure of concern shall also be implemented within the instrument examining the users self-assessed level of concern. An example of such is found within literature from various other sources. For example, such a question may take the following form:

1. “In general, how often are you concerned about your privacy while you are using the internet?” (Dwyer, Hiltz et al. 2007).

Westin derived a similar question where the possible responses required the participant to declare themselves as: very concerned, somewhat concerned, not very concerned and not

concerned at all (Westin, Maurici et al. 1998). Hence, a similar question and response is adapted and utilised in this instrument and takes the following form:

1. How concerned are you about your personal privacy when using a social networking site?

This is combined with another Westin oriented question ascertaining whether the user has been a victim of a privacy invasion and shall be present in the survey as follows:

1. Have you ever been the victim of a perceived invasion of your personal privacy?

Participants shall respond with options: yes, someone I know, no and do not know.

Finally, for measuring concern, it has been noted that thinking about privacy and becoming aware of certain features of social networking sites can raise participant concern and possibly promote a change in action (Pötzsch 2009, Tuunainen, Pitkanen et al. 2009). Therefore, a final question segment of the survey shall test if new knowledge raises concern and if this shall produce a self-reported intention to change behaviour. This shall have the obvious benefit of providing some indication of the potential effect of awareness on behaviour which informs and justifies the second portion of this thesis dealing with behavioural experiments.

As in the Tuunainen, *et al* (2009) paper the questions shall be focussed on a piece of Facebook's privacy policy dealing with the ownership of data and shall be posed as follows:

1. Are you aware that Facebook owns any information uploaded into the site (i.e. are you aware that it owns your information)?
2. If yes, does this affect your behaviour on the site (i.e. are you less likely to put certain information on there)?
3. If no, will you now modify your behaviour (i.e. are you less likely to put certain information on there)?

The goal of the above is to examine how aware participants are of some pressing and contentious privacy concerns surrounding the use and practices of Facebook. If they are aware, does that affect their perceived behaviour? This is to explore the idea that users behaviour is informed by their knowledge and their view of that knowledge.

Following this, is a final measure of concern reiterating the self-reported question described earlier; the aim here is to ascertain whether concern has increased over the course of the survey or in the face of new knowledge. These measures of concern shall be interspaced throughout the survey instrument and the complexity of privacy and concern shall be explored through the variety of measures implemented.

4.1.2 Measuring Attitude

Recall, an element of the paradox, as explored within past research, included measures of attitude; that is, how participants view the use of social networks by examining of usage patterns (Acquisti and Gross 2006). Hence, for a holistic view of the paradox and a complete picture of end-users privacy perceptions this instrument is required to include measures of social network use. Again, this keeps the work in line within the current research field providing greater validity of data.

This work sees attitude as a participant's use of social networks in terms of the extent to which they are active within the network. As such, typical questions here would include ascertaining if they are members of the network, how often they use the system and how many other users are in their friends lists. The style and types of questions implemented within this instrument shall be similar to those found within Jones and Soltren (2005), where a comprehensive survey was conducted to ascertain the threats present within Facebook. The survey included questions relating to who they add as friends, familiarity with the privacy policy, how many friends the participant has and how often they utilise the service. This specific research was not conducted in relation to the privacy paradox; however, the style and type of questions can easily be adopted and, indeed, are relevant to this research also.

A similar line of questioning was used by Strater (2007) which did deal specifically with the paradox; here, the focus of the study was disclosure within social networks. Furthermore, similar questions have also been implemented to measure awareness of privacy issues within social networks (Govani and Pashley 2005), where awareness was measured against the attitude questions to be implemented here. Hence, a complete picture of the paradox would not be possible without the application of a similar line of questioning. Therefore, the following table summarises the questions and responses found within literature and to be implemented in this study's instrument:

Table 2 – A summary of questions and responses

Questions	Responses
How regularly do you use a Social Networking Site?	A. Many times a day B. Once a day C. Many times a week D. Less than once a week
Have you read the privacy policy related to the Social Network System?	A. Yes B. No
Why do you use it (tick as many as apply)	A. Keep in touch with friends B. Keep in touch with colleagues C. Get to know new people D. Easily obtain information regarding work/university E. Show information about myself/advertise F. Make it convenient for people to get in touch with me G. Build relationships H. Find Jobs I. Other, please specify;
How many friends do you have listed in the Social Network System	A. 0-50 B. 50-100 C. 101-200 D. 201-400 E. 400+
What type of people do you add as a friend on Social Networks? (tick as many as apply)	A. Close friends B. Family members C. Friends you may not consider close D. Colleagues you may not consider friends E. People you know but do not consider friends F. People you have met but once G. People you have never met H. Other, please specify;
Do you use the "custom" feature to group you friends list into types of people?	A. Yes B. No
If no, why not?	A. Unaware of ability to do so B. Aware but do not know how C. Do not want to utilize feature D. Too time consuming to do so E. Other, please specify;
What would a person not on your friends list be able to see do you believe?	A. My Friends B. My Groups/Networks C. My Info D. My Pages E. My Photos F. My Wall G. Do not know

The previous table details the questions to be implemented in order to study the attitude of social network users. Attitude is defined here as their view of social networks which is indicated by their general use (i.e. not their specific behaviours) if social networks; so, how often do they use, how do they use it, do they read the privacy policy. The assumption from a privacy paradox point of view is that increased concern should be linked with having a lower number of friends or using the service less than others. Such an assumption is, in part, what this research is exploring by using a wider view of the phenomena.

One final question is added to this section which aims to build on the examination of the concept of privacy and inform future efforts to define the concept from a user point of view. This is an open-ended question asking the user for their definition of the concept of privacy summed up into one sentence. Such a question is introduced by this work to examine how the end-user views personal privacy. This question shall be posed thus:

1. Please write a brief sentence on what you believe privacy means to you.

It is left deliberately open in order to allow user to explain what privacy as they see it and how they relate to it; i.e. either generally or specifically to social networks.

4.1.3 Measuring Behaviour

The final two measures required by the survey concern behaviour. Specifically, a measure is required for self-reported behaviour in terms of privacy settings and self-disclosure of personal information (the second measure required shall be discussed in greater detail in the approach section).

Recall, there is a need for the study of behaviour to be granular in nature as behaviour within social networks is performed and protected individually; for example, a user date of birth is entered into a field on its own and separately from other pieces of data. Such aggregation of data is how the "digital person" is constructed within these social network environments (Solove 2009) and users are therefore required to deal with their personal privacy individually and granularly. Furthermore, granular information control is how the observable elements of the paradox are exhibited; i.e. an observed behaviour is either a piece of disclose information or the application of protection to individual pieces of data or groupings of data.

Stutzman (2006), talks about an issue with social networks being that new types of data, not normally thought of as private, are introduced requiring individual control and when taken

into account with various other pieces of information produce privacy problems. Hence the need for a granular study is suggested. Starting with disclosure behaviour; within the Stutzman study participants were asked to declare if they disclose individual pieces of information about themselves within social networks. A list of typical pieces of information found within services such as Facebook was generated and used as the question set within this study and this instrument shall adopt a similar approach. Therefore, within the survey to be used in this study the typical granular pieces of information found within Facebook shall be used as a complete question set examining the likelihood of disclosure as perceived by end-users.

However, differently from other studies this instrument intends to combine this exploration with the complexity of social privacy (continuing the theme of privacy complexity found throughout this instrument). For example, the role of conflicting social spheres has been noted previously (Binder, Howes et al. 2009) as a problem within social network as users tend to not only add their close friends to their network (Gross and Acquisti 2005). In order to encompass this into the survey participants shall be asked to measure their likelihood of disclosure to certain groups of people; namely, friends, colleagues and strangers. The following is an example of questions and responses which shall constitute the final instrument:

Full Name	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Date of Birth	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Note, the use of a Likert scale to for participant to grade their likelihood of disclosure; this carries benefit of allowing users to grade their perception giving a finer grain of responses.

Finally, for self-reported behaviour is the application of privacy settings. As in the Tuekci (2008) paper exploring the paradox, participants shall be asked to report on their privacy settings within social networks. In keeping with the idea of granularity participants shall be asked to report on what they believe their settings to be in a set of question derived from the granular controls present within Facebook. The response to such questions shall be based on the options available to be applied found within the system as the following example of the instrument shows:

Primary Information	Everyone	Friends and Networks	FOAF	Only Friends	Custom	Not Sure
Full Name						
Date of Birth						

Note these granular groupings, in terms of their presence within Facebook, were true at the time of the surveys administration to the sample group.

4.1.4 Observing Behaviour

The final requirement of this phase of the study is to gather observed results from actual network use in order to ascertain whether the paradox is still in evidence within the sample available to this research as a whole. Therefore, a portion of the sample must be analysed for their actual behaviour within the network. Using the same table for reported behaviour (to maintain consistency) a sample of the participants who responded to the survey shall be examined in person by the researcher. If there is data present and available to the review of their profile then it is said to be disclosed and open to strangers; this can then be compared to the reported answers given within the survey. So, a Facebook profile shall be used which holds no relation to the sample taken. If a piece of information is visible on their profile when visited, say their date of birth, it is present to strangers. This then is compared to whether they reported as such in the survey. As the information on a profile falls into several categories of settings it is therefore possible to determine whether their profile is as open as they believe. Any information governed by a setting other than “open to Everyone” should not be visible to the profile examining the samples profiles.

4.2.1 Detailing the Approach

The research will be conducted within the University of Salford and shall focus on first year undergraduate students for a number of reasons. First, the paradox has been noted as being a particularly strong feature of young users of social networks (Barnes 2006) within the age range of a typical undergraduate. Secondly, the research shall use convenience sampling and approach participants within lecture sessions as undergraduate classes will offer larger numbers of potential participants than other methods. Convenience sampling is utilised due to the researcher's access to student within the university, this will allow maximisation of potential participant numbers.

Therefore, lecturers in charge of modules with large student numbers shall be approached via email to request their cooperation with the research. Those who respond shall arrange a date and time for when the research can take place within one of their sessions; either before lecturing or after. The survey shall be administered as agreed upon alongside consent forms for participation in the research. Those participants who indicate a willingness for further involvement in the work shall be used for an expert evaluation of their social network profiles in order to ascertain actual behaviours. These evaluations will be conducted immediately after the survey administration to ensure a current (in terms of data collection) snapshot of user perception and behaviour.

The reported settings from users shall be collated and turned into a "privacy score" and the observed settings given the same treatment. This shall allow for the scores to be compared using statistical tests giving greater validity to the results set. Scores are generated thus: a point of 1 is added if the information grouping is reported to be available to strangers and if it is available to strangers when reviewed. Hence, a higher score would indicate an increased amount of available information either reported or observed. For example, Facebook's default privacy settings (at the time of data collection) would have a score of 4; name, photos, basic information and friends lists are available by default. Incidentally, this score shall be used when analysing data to ascertain deviation from the default.

The full survey instrument described thus far can be found in appendix 2.

4.2.2 Initial Limitations

The sampling method chosen is not without limitations. First, focussing on a specific population for the sampling method may not be representative of social network users as a

whole. However, it is noted that the paradox is a phenomenon typically related to teen users of the service; therefore, such a focus is entirely appropriate. It is also noted that the response to the email may not be all encompassing of the student population present in Salford. For example, a survey including only those students on technical based courses, such as computer science, may not present paradoxical privacy behaviour which is representative of a SNS population as responders may be more technically aware and privacy savvy than other users. The results therefore, could suffer from a lack of generality to the wider research field. This is an acknowledged limitation of the approach and steps shall be taken to ensure a range of courses are involved in the research to enable a range of participants. Specifically, when approaching potential participants, courses from across the university shall be targeted.

Finally, the limitations of a survey method have been noted in the philosophy section; however, it is worth reiterating the need for such an approach in this work. The main concern here is to quantify and measure the aspects of the paradox in order to empirically explore the relationships which exist within it (either finding causal relationships or not). Furthermore, consistency must be maintained with the theme of the study as a whole which is exploring observable and quantifiable behaviour through HCI experiments.

4.2.3 Summary

This chapter has described the creation of an approach to studying the privacy paradox phenomenon in a way that is more complete than any implemented within research thus far. A survey instrument is developed that measures each aspect of the paradox identified, but not completely studied, by literature; concern, network use and reported behaviour. The approach also details a method of behavioural observation to allow for comparisons to the survey responses; hence, all appropriate aspects of the phenomenon can be recorded and examined empirically. A sample shall be obtained for the study from the undergraduate first year population within the University of Salford following a convenience sampling method.

Chapter 5 – Survey Results

Chapter 5 - Privacy Perceptions Survey

5.1 Introduction

Figure 10 from the literature review chapter provides an outline of the relationships being explored through this survey. Specifically, literature often assumes there should be a correlation between concern and attitude and between concern and behaviour (higher concern the less likely users are to disclose information and to apply protection). Literature also assumes that intention should match actual behaviour. This survey shall therefore test the following statements using the measures outline in the design chapter previously:

1. Measures of concern are associated to the level of protection reported and applied within Facebook.

2. The settings reported by users match those that they apply within a real world social network.

The survey is aiming to quantify the privacy paradox as it exists within sample (if it does) and clarify where the paradox potentially takes place and to examine if varying measures of concern give rise to varying relationships. Furthermore, the survey method is explored as an appropriate tool for exploring the concept of privacy and the paradox.

As per the research methods outlined, the survey was conducted prior to lectures during teaching time with the assistance of the lecturer leading the session. Out of around 400 potential participants, 340 responses were obtained from the following modules; English literature, Spanish, French, Research Methods (Nursing), Analytical and Research Skills (Law) and Foundation Module (Nursing). A response rate of around 85% was, therefore, gained for this survey. Furthermore, 151 of the survey responses where checked for actual privacy behaviour within the social networking system Facebook, this accounted for 43% of the participants who took part in the survey.

5.2 Participant Overview

This initial section shall briefly go over the participant details for the sample used in this study (full details can be found in appendix 3). Mentioned in the introduction, the full number of responses was 340 with the following table breaking down the gender split;

Table 3 – Gender Breakdown

		Frequency	Percent	Valid Percent
Valid	Male	70	19.9	20.6
	Female	270	76.9	79.4
	Total	340	96.9	100.0
Missing	System	11	3.1	
Total		351	100.0	

The number of females taking part in the survey far outweighs the number of males; this could be due to the convenience sampling method used where the courses who responded to the request for participants mainly came from the nursing school which predominantly is female oriented within Salford. This is a noted limitation of this sampling method.

Participants were also profiled according to their age and can be broken down as such;

Table 4 – Age Classification

		Frequency	Percent	Valid Percent
Valid	Non M'	201	57.3	59.1
	Mature	139	39.6	40.9
	Total	340	96.9	100.0
Missing	System	11	3.1	
Total		351	100.0	

As per the University classification system, a participant is classed as mature if they are over the age of 21; which shows a 60/40 split in the participant group.

The following table shows the breakdown of participants who reported that they are users of social networking systems;

Table 5 – Social Network Users

		Frequency	Percent	Valid Percent
Valid	Yes	301	85.8	88.5
	No	39	11.1	11.5
	Total	340	96.9	100.0
Missing	System	11	3.1	
Total		351	100.0	

As can be seen, the number of users in the sample far outweighs the non-users showing the prevalence of social network today. To further make this point the following table shows just how often such systems are used, with nearly 80% of the sample reporting at least daily visits.

Table 6 – Regularity of Use

		Frequency	Percent	Valid Percent
Valid	Many a day	139	39.6	46.3
	Once a day	92	26.2	30.7
	Many a week	52	14.8	17.3
	Less one a week	17	4.8	5.7
	Total	300	85.5	100.0
Missing	System	51	14.5	
Total		351	100.0	

The final table to be considered when examining the sample details the number of reported friends participants reported as having in the social network.

This table is interesting as the average number of Facebook friends (in 2011, when this survey was conducted) was around 130 (Quercia, Lambiotte et al. 2012). As shown in this study the most selected grouping of participants was in the 200-400 range, which is considerably higher than the average. Again, this could be put down to the convenience

sample and the fact that this sample is from only a particular segment of Facebook's demographic. However, this does match with statistics from the Pew Research Institute where the average was 226 friends (Hampton, Goulet et al. 2011). Furthermore, research shows that younger users (of which this sample comes from) tend to have much larger friends lists than other demographics (Ugander, Karrer et al. 2011).

Table 7 – Reported numbers of friends

		Frequency	Percent	Valid Percent
Valid	0-50	13	3.7	4.3
	50-100	27	7.7	9.0
	100-200	74	21.1	24.7
	200-400	98	27.9	32.8
	400+	87	24.8	29.1
	Total	299	85.2	100.0
Missing	System	52	14.8	
Total		351	100.0	

5.2.1 Detailing Concern

The survey contained three separate measures of concern: the Westin rating, a simpler self-classification measure and a similarly worded question at the survey's conclusion. The goal of using three approaches to measuring concern was to examine the complexity of it from a user perspective; would there be a consistent measure throughout? Furthermore, a greater chance of causal relationships would be possible through more than one measure; i.e. if one does not relate to behaviour then would another?

First, then, the Westin rating (which assigned a rating based on the response to a set of three questions) is detailed in the following table:

Table 8 – Westin Spread

		Frequency	Percent	Valid Percent
Valid	**Unconcerned**	60	17.1	18.1
	Pragmatist	195	55.6	58.9
	Fundamentalist	76	21.7	23.0
	Total	331	94.3	100.0
Missing	System	20	5.7	
Total		351	100.0	

When conducted in wider research, the Westin rating typically is split across sample in the following way: 25% Fundamentalists, 57% Pragmatists and 18% are Unconcerned (Kumaraguru and Cranor 2005). Studying the above results shows that they fit very well with the research field at large matching the percentages very closely; in terms of measuring this degree of concern, therefore, the research has been shown to have validity through generality of results.

The following measure of concern came immediately after this one on the survey and was also designed by Westin. It is, however, a much simpler rating of privacy allowing the participant to declare for themselves whether or not they are worried about their privacy in social networks.

Table 9 – Concern Spread

		Frequency	Percent	Valid Percent
Valid	**Concern**	206	58.7	62.6
	No Concern	123	35.0	37.4
	Total	329	93.7	100.0
Missing	System	22	6.3	
Total		351	100.0	

Interestingly the amount of self-reported no concern is higher than the Westin rating of Unconcerned; illustrating the variance that different method of measuring concern can introduce. However, the two do hold an association according to a Chi-Square test ($x=33.7$, $p<0.0001$). Figure 11 illustrates this.

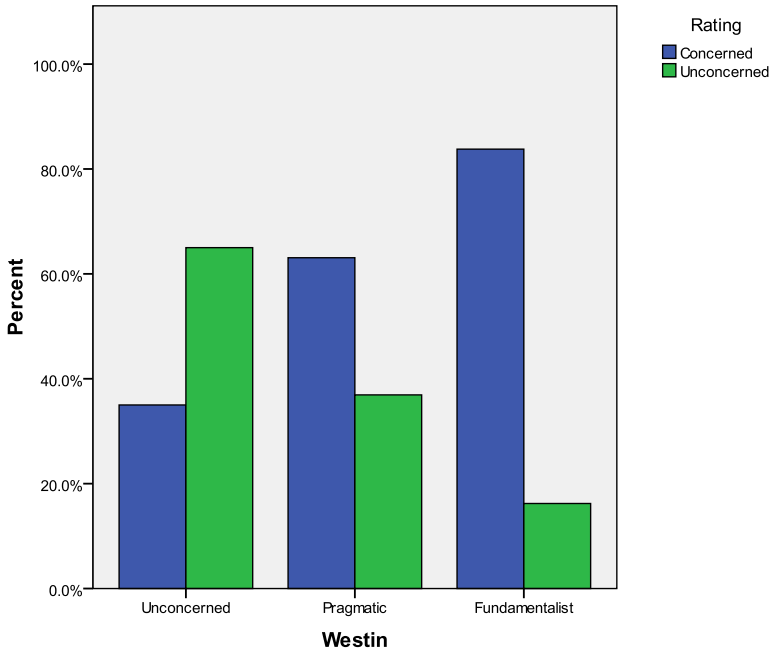


Figure 11 – Westin – Concern Relationship

Figure 11 suggests that the Westin Unconcerned (on the x-axis) have a much higher degree of self-reported unconcerned (represented by the green bars) people and the opposite is true

for the other Westin ratings. While this result is unsurprising it does show that the two measures are related.

Notice the green bar in the Fundamentalist category showing users who reported that they are unconcerned about social network use. Although only 12 participants fell into this paradoxical category, it nevertheless shows the problem with concern as a measurement or influence of SNS use. For example, of the 12, 7 had the lowest score for their reported privacy settings (i.e. the strongest level of protection) which is perhaps to be expected going by the Westin Index. Their actual scores (of those involved in further study) were actually higher than their reported, suggesting the privacy paradox at work in some way.

A similar analysis of the Pragmatist rated participants better illustrates this point with little difference between the reported scores and the actual scores of the self-described participants in this rating.

First, the reported scores (scores are marked out of 10 and bracketed in a score out 5 represented by the RepPBracket). A low score indicates a high level of wanted settings and a high score a low level.

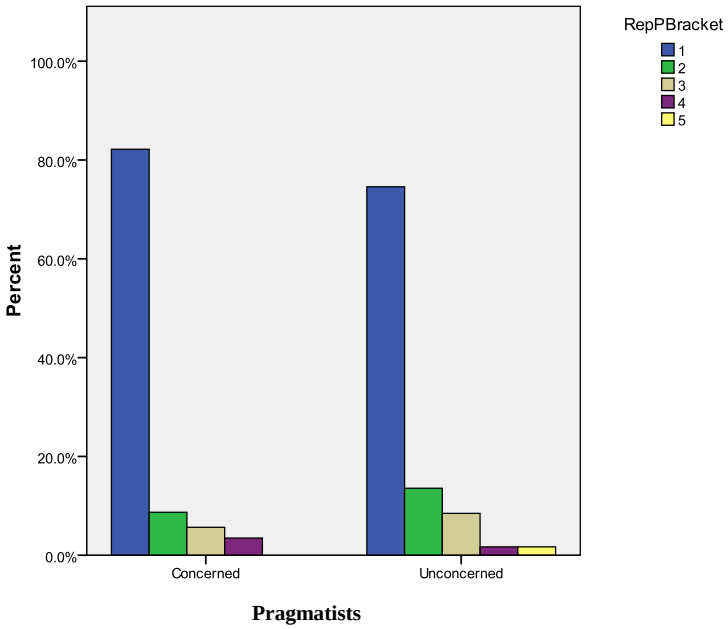


Figure 12 – Reported settings – Self-described concerned and unconcerned pragmatists

The graph is showing the reported scores for the Pragmatist rated participants split into self-reported concerned and unconcerned. Note, that the spread across the two groups is very similar in terms of what settings participants believe they are applying regardless of their level

of concern (a score of 1 would be achieved through everything being available to friends only). Indeed, there is no statistically significant difference between the two as may be expected (Mann Whitney U p=0. 84).

The graph for actual, observed, scores is as follows:

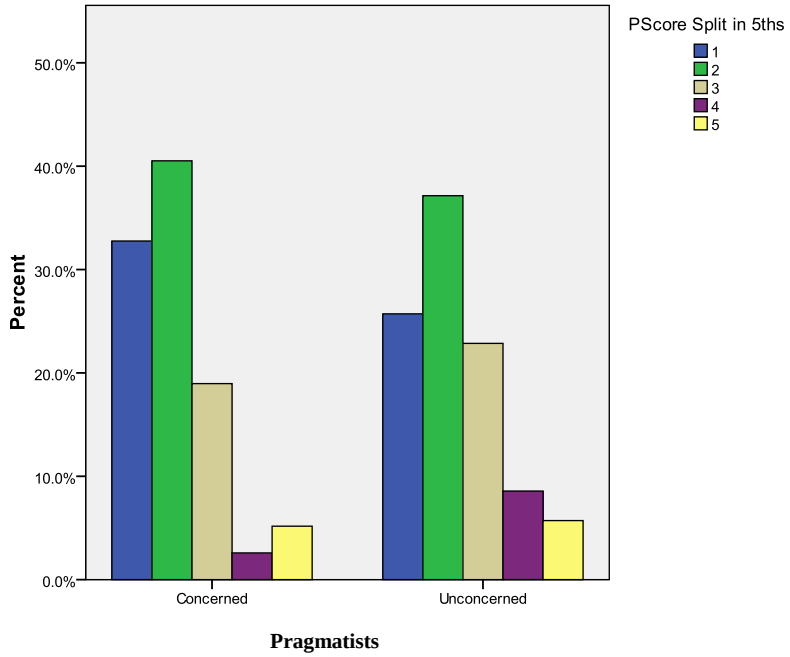


Figure 13 – PScores for Pragmatists

Again, there is a very similar spread with no statistically significant difference (Mann Whitney U p=0. 342), although the unconcerned group had some scores which are slightly higher, this was not significantly the case.

So, what does this imply? By studying the pragmatists in the survey the two graphs show that, regardless of concern, participants have very similar levels of reported behaviour and actual behaviour where their privacy settings scores are concerned. That is, regardless of their level of concern there is a consistent spread of how participants *want* to act and how they *do* act and their respective levels of concern play little part in influencing. One might expect, for example, that there be far higher "PScore" in the unconcerned groupings of participants. This has been a slight deviation from the chapter flow and shall be picked up again later, first, however, what else can be inferred from reported concern?

5.2.2 Reported Concern and Social Network Use

Given that concern forms a part of the privacy paradox according to literature, each of these above measures shall be examined for relationships between themselves, general network use and reported behaviour.

The Westin measure of concern held a Chi-Square association with the number of friends participants reported ($x=21.119$, $p=.007$); similarly, the self-reported level also did ($x=11.324$, $p=.023$). Interestingly, Westin rated Unconcerned participants held fewer levels of the top-tier number of friends (400+) than the other two groups; this could imply that participants are occasionally unconcerned about their privacy if they are confident in their ability to protect it; i.e. they know that they have a restrictive and trusted friends list and as such are unconcerned about the risk. More work would be required in order to research this potential phenomenon further; however, it is not unreasonable to assume that confident, skilled user would be unconcerned about their privacy.

This theme is, in some way, continued in the next relationship where pragmatists and fundamentalists are more likely to report that they are unsure of their privacy settings; although, this is with a p value of .05 so may not be a statistically significant relationship ($x=5.97$, $p=.05$). This is similarly true in regards to the self-reported measure where concerned participants are more likely to admit to being unsure of their settings ($x=5.48$, $p=.019$). This could lead to the following conclusion, that concerned participants are aware of their own limitations and this informs their level of concern. Hence, the flexibility of concern is by no means a definitive indicator of behaviour as participants could, feasibly, be concerned due to their own knowledge and awareness of privacy and, also, have lax privacy settings due to this.

The Westin concern rating held no other relationships with reported use of social networks for the sample as a whole. Self-reported concern, however, had a few more. First, concerned participants are more likely to have read the policy of Facebook ($x=10.04$, $p=.002$) and are slightly more likely to report implementing custom privacy settings ($x=4.09$, $p=.043$). Each of these relationships makes sense if it is assumed that concern is a valid indicator of a more careful privacy outlook of social network use.

This brief section has explored concern and its relationship to a variety of simple social network usage indicators. It is by no means exhaustive but does highlight some interesting

points regarding the idea of concern. First, concern itself, like the concept of privacy, is changeable and means different things to different people. One participant's concern could be low as their skill level is high leading to limited disclosure. In this instance concern is not an indicator of behaviour without knowing more about the participant in question. Secondly, how concern is measured is indicative of the relationships it could form with other elements in social network usage; some measurements hold relationships, while some do not. Hence, due to the nature of concern it is perhaps unreasonable to assume that clear indications of behaviour can be gained from any one definition concern alone.

This is further exemplified through the final concern measurement which deals with ownership of data. The response is classified in the same way as the self-reported measure of concern previously dealt with and, as it deals with an element of Social Networks, concern of it will indicate a concern while using the system. Hence, an initial report of unconcerned and a final measure of concerned could indicate a change in that participant's perception of social network use.

First, then, the results themselves; as can be seen from the following table there is a statistically significant (Wilcoxon Signed Ranks $p<.0001$) increase in the number of participants reporting concern compared to the first measure at the start of the survey. Recall, table 9 detailing the initial, self-reported measure of concern and compare it to the second at the end of the survey in table 10.

Table 10 – Self-Reported Concern, Measure 2

		Frequency	Percent	Valid Percent
Valid	Concerned	217	61.8	76.7
	Unconcerned	66	18.8	23.3
	Total	283	80.6	100.0
Missing	System	68	19.4	
Total		351	100.0	

Wilcoxon is used as it is a repeated measure using the same participant group (Ott, Longnecker et al. 2001). Therefore, a significant number of people became concerned about

an aspect of social networking when they became aware of new information (specifically that Facebook owns the data stored on it). In total 74 participants who were initially completely unconcerned about their privacy in social networks expressed a final concern regarding this point (non SNS users did not answer this question, hence the decreased total). So concern is changeable depending on how the question is being asked and on the, possibly changing, perception of the question from users.

5.3.1 The Privacy Paradox

In terms of Figure 10 the previous section has preliminarily explored the element of concern as an indicator of general social network use which may be considered an aspect of user behaviour. However, disclosure and settings are the defined, observable behaviours which have been defined as discrete manifestations of privacy behaviour in social networks. This section shall now deal with examining if users are acting as they say they are (link between intention and behaviour).

Profiles were analysed according to the presence of data from 10 pre-defined groups representing the privacy settings found within Facebook. A score of 1 was added to an overall score if one of these groups was open to everyone visiting the profile and the final score divided by 2 to provide a manageable spread. For example, Facebook's default settings have a score of 4 (the following groups are open to everyone: wall posts, friends list, photos, and basic information) which is bracketed to 2. The reported settings scores are detailed figure 14.

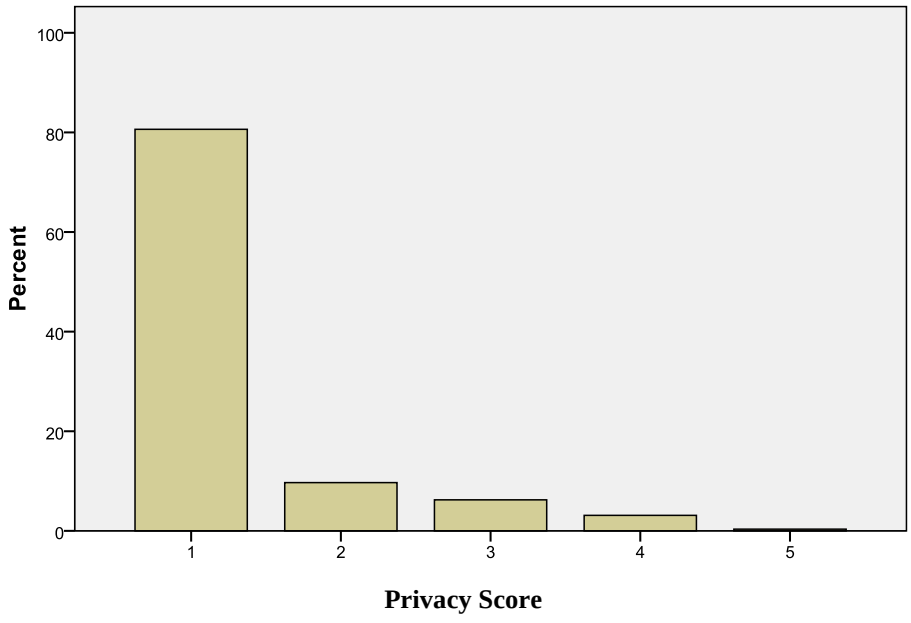


Figure 14 – Reported Scores spread

The graph shows that the reported scores of 80% of the participants are in the lowest bracket. Indeed, as can be seen in the table found in appendix 3, the modal score across the sample is 4 for all settings except full name. That is, the most select setting for all information groups was "friends only"; therefore, the majority of participants felt that their actual information was open only to friends except for their full name (so profile can be found but not viewed).

From these results it can be seen that the majority of participants in this sample believe their privacy settings to be fairly high, that a stranger idly viewing profiles would only be able to see their full name (and only 36% of the sample chose "Everyone" in this category; so, while modal, it is not as large an amount as other information groups).

Comparing these results to the settings scores gained from actual observations reveals evidence for the privacy paradox:

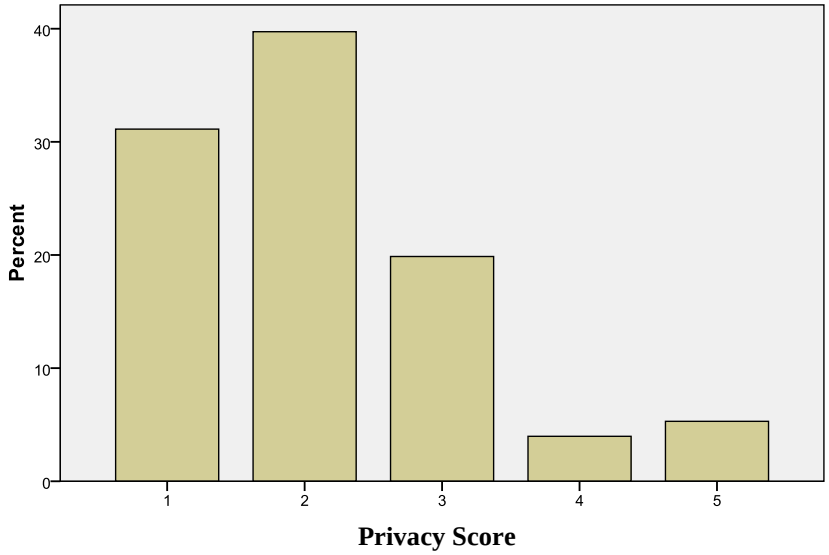


Figure 15 – Observed Scores (PScores)

Immediately, it is clear that the scores are, on average across the sample, increased from the reported scores. This is a statistically significant increase with a Wilcoxon score of $P < .0001$; it is, therefore, a highly significant increase from the reported behaviours. Note the most common score here, with around 40% of the sample, is in the 2 bracket. Recall, that the bracketed score for Facebook's default settings was 2; this would suggest that participants are simply leaving the system in its default state. Indeed, this is in keeping with the finding of wider research where it is estimated that around 80% of users in social networks leave their settings at default (Bonneau, Anderson et al. 2009). This work suggests that the figure is not

quite so high yet does provide evidence that leaving the settings at default is an issue within this sample.

Therefore, point two in the introduction is not true, as the reported settings scores and the actual scores do not match, but in fact increase with statistical significance providing evidence for the paradox.

5.3.2 Specific Cases

In order to fully appreciate this point in greater detail, a few specific cases shall be examined for the increases they hold.

First, participant ID#17 seems to demonstrate the hallmarks of the privacy paradox evident in end-users. A reported score of 1 does not match with the actual, observed score of 3 shows that the participants profile and information contained within it are actually fairly open. Furthermore, it is actually higher than the default suggesting that the participant changed the settings to a more open level which is not reflected in the reported levels. Examining concern further clouds the issue; with a pragmatist rating and a self-reported measure of concerned while using social networks, the reported scores are perhaps to be expected. These obviously, however, are in no way indicative of actual behaviour. Furthermore, the participant reported a friends list of greater than 400 and this is the case with up to 1700 friends listed in the observation; this not what would be expected from a concerned participant. However, the fact that this is reported accurately, while settings are not, is interesting.

Another interesting participant (ID#203) for further study exhibits somewhat different traits to the first. Similar to the above participants, a score of 1 was reported for their perceived privacy settings. However, the observed score is placed in the highest bracket of 5 showing a large increase from the perception. A reported level of unconcerned would explain the observed behaviour if using the usual assumptions. Although, when considering the reported scores the case makes less sense but can be (potentially) explained thusly; the participant's unconcerned attitude relates to the confidence from their perceived settings scores as this is based on a lack of awareness, this level of concern is ill-informed. Hence, the lack of skill and awareness of privacy is manifested in the loose privacy settings. This is further demonstrated through the final measure which records the user as concerned at the end of the survey.

Each of these two participants exhibit the same level of reported concern with an increased level of observed (actual) concern; however, their levels of reported concern are polar

opposites. This would demonstrate the pitfalls of assuming certain behaviours can be derived from any one measurement of concern. If examined across the sample some interesting results are found which are demonstrated and discussed in the following section.

5.3.3 Concern and Behaviour

The theme of this section was touched upon earlier with the examination of a sub-set of concerned participants and their related settings scores and shall be expanded upon here. To review the central finding from that section; there was consistent results across the sub-sets of concern studied in terms of the reported and observed behaviour. The aim here is to explore point one from the introduction based on the assumed link between concern, intention and behaviour illustrated in figure 10.

Dealing first with the Westin rating there is no association (chi-square) recorded with either reported settings scores or observed scores (x=2.67, p=.953 and x=6.06, p=.641); therefore, no significance at all for these two measurements. The following graph demonstrates this for the reported privacy settings:

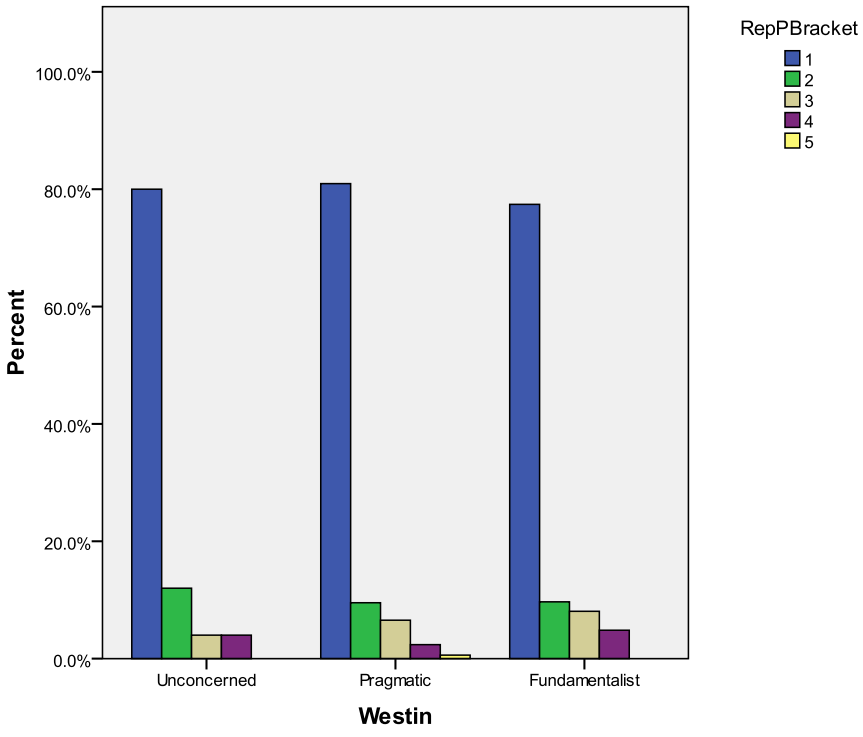


Figure 16 – Reported Scores across Westin

And for the observed scores:

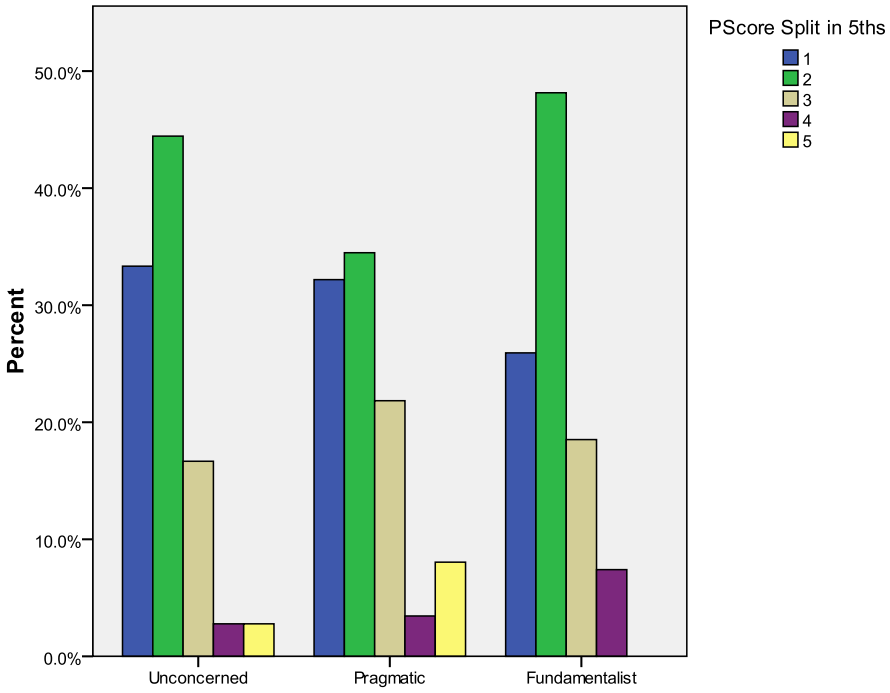


Figure 17 – Observed Scores across Westin

Notice, in the two graphs the spread across the three Westin groups is very similar. Although, in the observed scores chart the fundamentalist group has less no scores in the highest bracket, the difference is negligible and not statistically significant.

So, what can be summarised from these results? Regardless of the Westin rating given it has little to no bearing on a participant's likely scores for either of the two measures available (reported and actual). Also striking is the number of participants that feel their settings are "friends only" throughout the SNS profile, again, regardless of the level of concern they hold. It could be that having a belief in a high privacy setting would lead to a user being less concerned about their privacy as they believe it to be well protected. Or, to state the opposite (and assumed relationship) a participant who is highly concerned will have high privacy settings due to their worry. The potential flexibility of concern in these two cases could be a reason why the above two graphs are the case; that is, no relationship with either.

To further explore this, the self-reported measure of concern shall be subject to the same analysis. In terms of relationships the chi-squared test, again, found no statistically significant link between this measure of concern and reported settings or actual settings (x=6.86, p=.143 and x=3.59, p=.466 respectively).

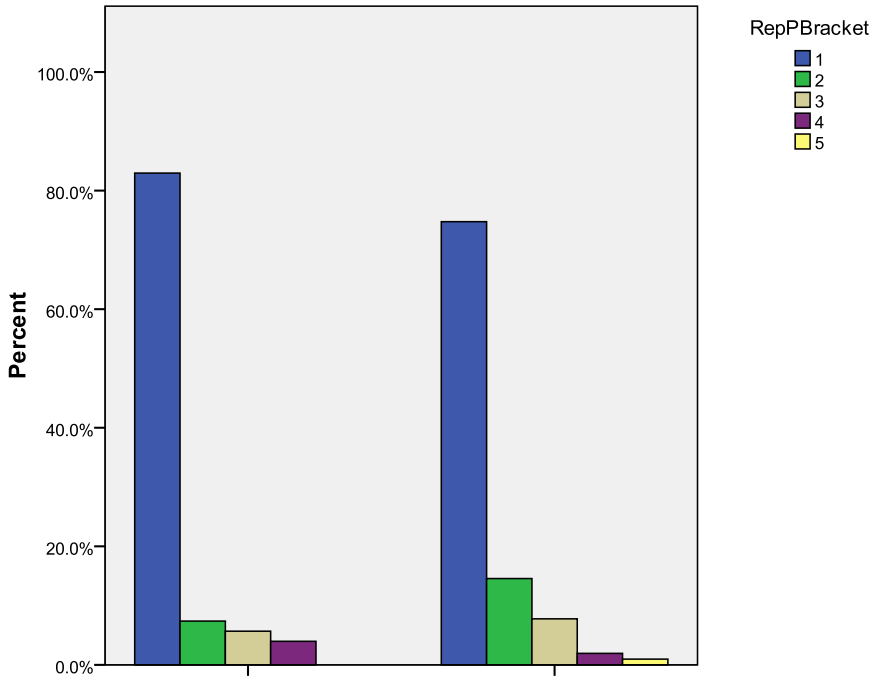


Figure 18 – Reported scores across concern

Initially, notice how the spread is very similar to the comparative chart in the Westin evaluation from earlier. The concerned scores are slightly lower than the unconcerned; as before, however, this is not statistically the case.

The following displays the spread for the observed scores:

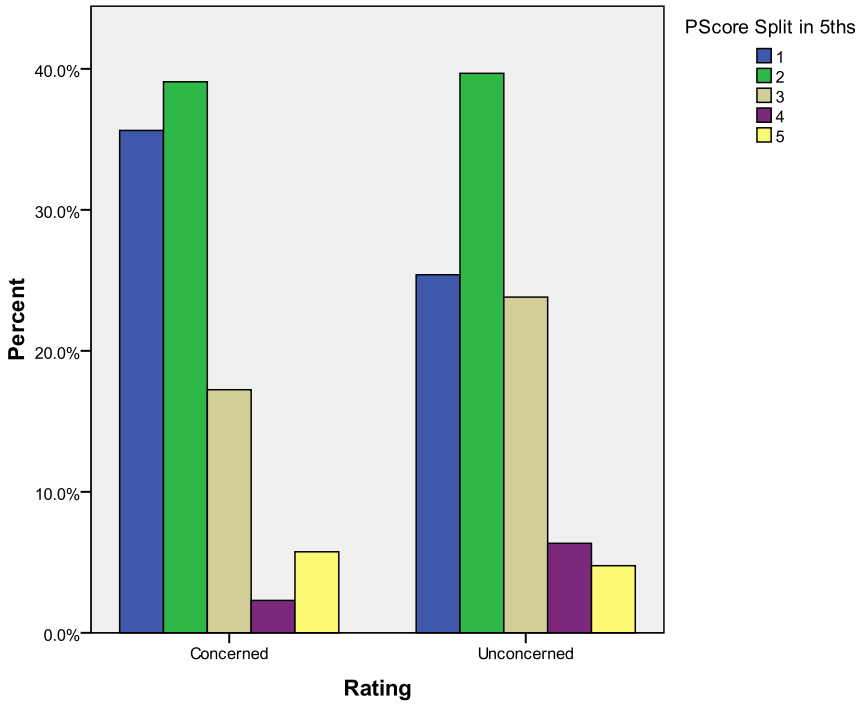


Figure 19 – Observed scores across concern

Notice, the scores do increase slightly for the unconcerned group with a greater amount of participants holding the middle bracket with their settings score; also, the concerned group held more participants in the higher bracket than unconcerned and this could be for some of the reasons outline earlier (low skill increases concern if the user is aware of it). Again these findings are without any statistical significance.

In order to further explore this point, two sub-sets of participants shall be compared; those who are concerned in both rating and those classed as unconcerned in the two ratings (essentially those participants who are very clear on their level of concern either way). In the first category (fundamentalist in Westin and concerned in self-reported) held 62 of the participants from the sample. Furthermore, the second category (Westin Unconcerned and Unconcerned in self-reported) held 39 participants from the sample. The following two graphs show the reported scores for the two polar opposite groups.

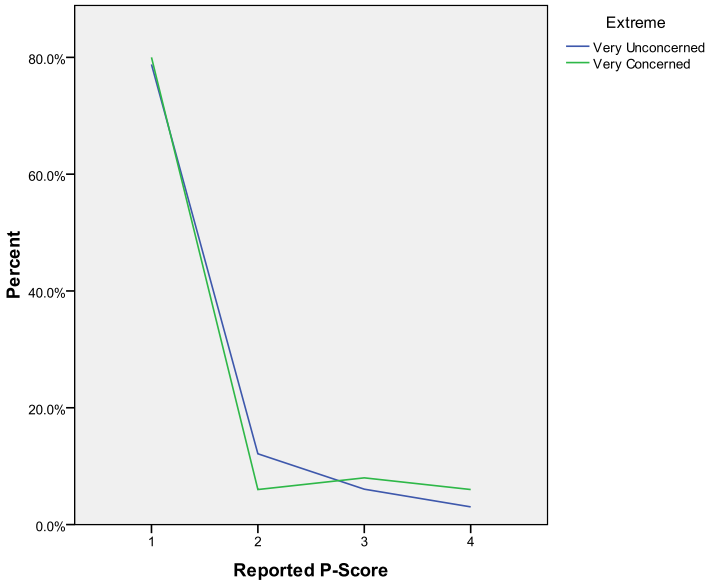


Figure 20 – Reported Scores – opposite groups

Notice that, even here, the two groups are very similar in their reported scores despite their very clear view on their level of concern. However, should this be an expected result? A concerned person is very likely to have a perceived or desired high settings score. Furthermore, an unconcerned person may very well be that way due to their perceived level of high protection.

Following this analysis further to the actual scores for the two groups gives the following graph:

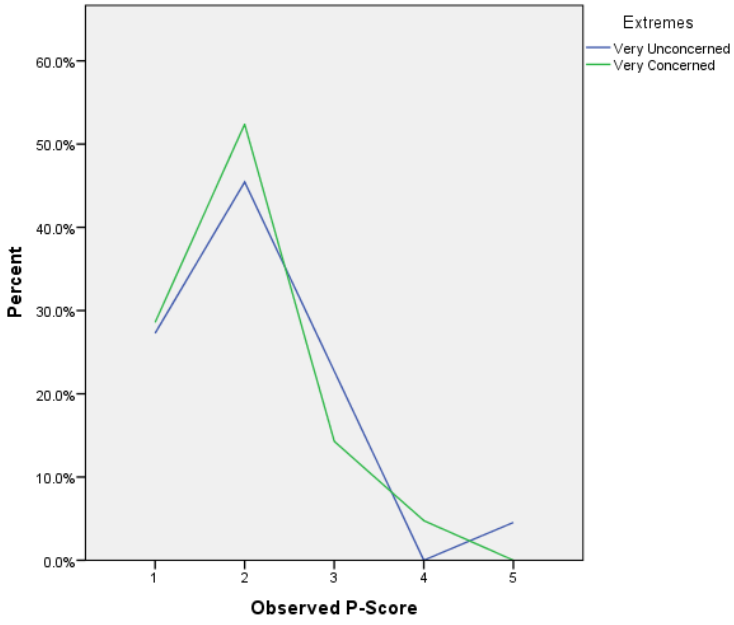


Figure 21 – Observed Scores – opposite groups

Again, the spread across the two groups is very similar and there are no significant differences between the two despite their clearly opposed ideas of privacy. It should perhaps be expected here that concern should play a role in how users *actually* behave and yet it clearly does not; with patterns of use being very similar across all measurments of concern. What is clear, however, is that the paradox is very much in evidence. Participants actual behaviour rarely matched the reported behaviour across the sample as a whole. Given that the effect is equal across all groups and measurements of concern, it is a fair assumption to make that a consistent entity is having an equal effect on all. That is, whatever is causing the paradox is enacting on all participants equally and, when talking about it, the essence of concern plays a smaller part in the phenomena than originally thought (within this sample).

The reported scores show that, regardless of concern, the majority of participants perceive their settings to be friends only (or want them to be) and therefore, the pieces of information they disclose on the social network is to their only friends. This could nullify any concern users have if there is confidence in the use of the system and the protection it has while unconcerned participants may very well be that way because they are confident.

Therefore, point 1 from the introduction is also not true as their appears to be no assocation between measures of concern and desired and observable behaviour as described by figure 10. However, there is a consistent desire for a certain level of protection acorss all measures of concern implemented.

5.4.1 Granular Privacy Perceptions

The questions dealing with likelihood of disclosure (measured on a Likert Scale of 1 to 5) to various social spheres are detailed within this section. This produced a large data set which is broadly represented by the following graph:

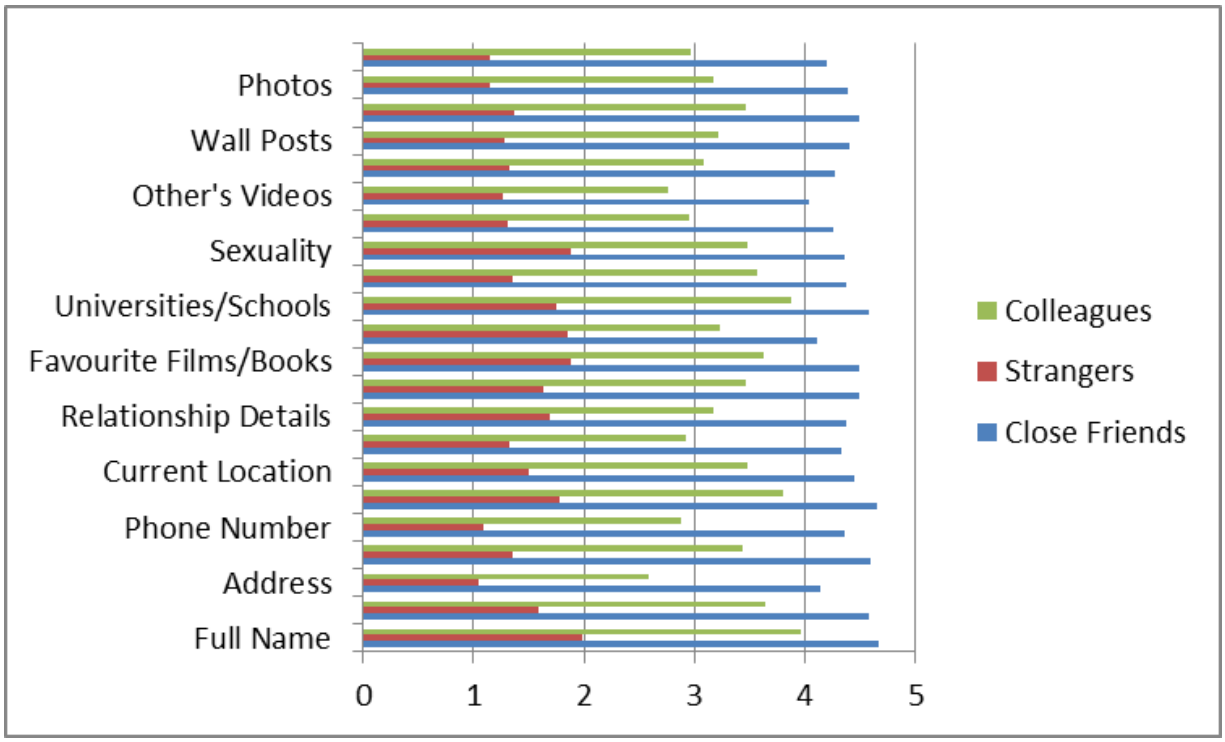


Figure 22 – Likelihood of disclosure to social spheres (5=high, 1=low)

From this graph the differences in social sphere perceptions is demonstrated. Participants rated themselves as very likely to disclose any granule of information to who they perceive as their close friends. Indeed, the modal selection for each of the granules in this category is 5 (very likely). This is in line with the perceived or desired reported P-Scores from the previous section where participants selected "Friends Only" for nearly all information groupings. The social sphere of strangers, however, is at the other end of scale with modal scores of 1 (Very Unlikely) for all data granules. Note the varying perception of sensitivity across the data items where participants rated physical data as more sensitive than more intangible ones. For example, address and phone number which grant physical access to the person where more highly unlikely to be disclosed to strangers according to the graph. This difference in the value placed on traditional pieces of information has also been highlighted within wider research (Stutzman 2006).

With regard to colleagues however, participants were far less certain in their ratings as a sample as a whole demonstrating the variation in the perception of trust placed on that social sphere. Some pieces of information trended towards the more likely to disclose (see Address) while others went the other way and are more spread across the possible responses (see Photos) showing that varying participants place varying value on the individual items.

A clear need for careful and personal privacy management is therefore required for the social spheres present within a social network. Although, participants specified their need for friends only settings they do not only add friends to their network as demonstrated in the following figure:

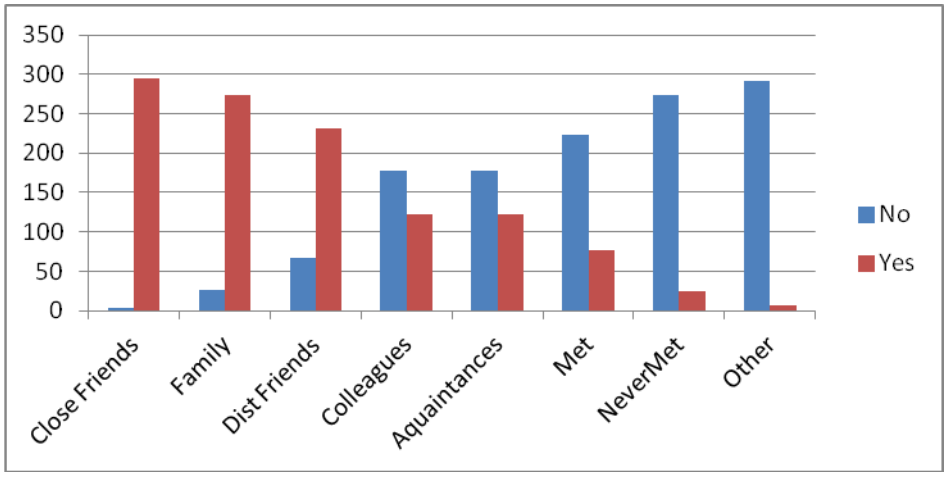


Figure 23 – Who people add to the lists

So, within the network there is a further need for social management with regards to the granularity of who can see what information when it is disclosed and in the future. The custom feature for example is one such tool that could be used to manage a user's social spheres so that only close friends from the "Friends List" see appropriate information. However, within this sample the reported use of the custom feature is low as demonstrated by the following figure:

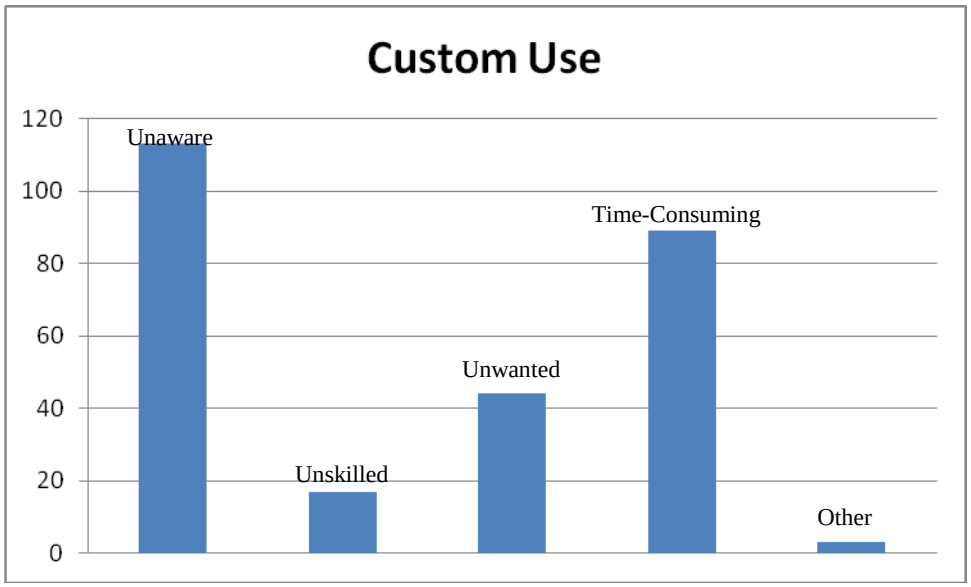


Figure 24 – Reasons for non-utilisation of Custom use

Furthermore, the results from the observed P-Scores show that participant's information is open to strangers as the researcher (who is a stranger in the network to participants) found information that they reported as being set to "Friends Only": a paradoxical disconnect between perception and action.

This section has demonstrated the complexity of conflicting social spheres and the underutilisation of tools within the network to manage those social spheres. Participants do desire to disclose to only certain parties as reflected by their reported likelihood to disclose and matched by their desired reported P-Scores (settings). However, their action in the network does not match, perhaps because that complexity is too difficult to manage consistently through changing social contexts. Also, the results show individuality of the privacy problem adding another layer of complexity. The colleague's social sphere, for example, had much more spread responses across the data sensitivity questions showing that the perception of trust in colleagues is different from participant to participant. This aspect of individuality is worth noting further as it demonstrates the need for users to meet their own privacy needs based on their perceptions. In short, a technological solution could be considered difficult as management of this individuality is only possible by the individual.

Other examples of this individuality were in evidence throughout. For example, Westin fundamentalists were more likely to report having fewer friends than their counterparts (Chi-Squared $x=21.1$, $p=.007$). There also existed a significant relationship ($x=6.20$, $p=.013$) between the self-reported rating of concern and gender were females responded as being more concerned of their privacy in social networks (figure 25).

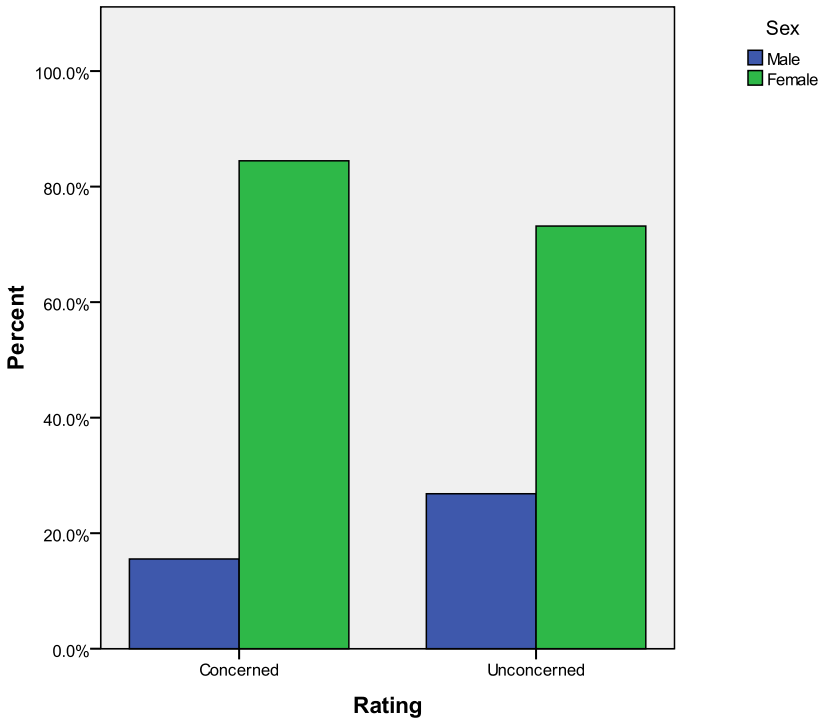


Figure 25 – Gender and concern

This, therefore, is consistent with what would be expected given the conclusion drawn from the above graph; females in the sample have higher levels of reported concern generally and this is reflected in the responses to certain data granules. Indeed, wider research has found similar findings (Fogel and Nehmad 2008).

So, why demonstrate individuality? The responses to these questions have shown that, where certain social spheres are concerned, participants have individual perceptions of privacy some of which are related to the kind of person they are. This, then, shows privacy to be dependent on the individual and, if their needs are to be met, then a granular level of control is required. However, it must be one which is implemented by the individual or they cannot be catered for otherwise. Participants do have nuanced views of appropriate disclosure which is personal to them and their reported behaviours and corresponding perceptions are consistent with each other. However, the consistency breaks down within the social network itself which begs the question of why; how are participants persuaded to veer away from their reported, consistent beliefs?

What can further be taken away from this? Participants have a clear definition of who they wish to disclose information to and that shifts and changes depending on what that piece of data is and who the party is learning of that data. Furthermore, the decision to disclose is also

dependant on who is doing the disclosing; although there are clear trends across the sample as to the general idea of the likelihood of disclosure to certain groups, there exists a more even spread where colleagues are concerned and further demonstrated through the relationships explored between groups and the data granules. Hence, the individuality of privacy is also demonstrated from the results here. Despite this, the custom tool to deal with this complexity is not utilised by the majority of participants for a variety of reported reasons. The main reported reason was one of a lack of awareness; the question, then, is what would happen if users were made aware of the custom tool? Or, what would happen if it was made easier to use? Would either of these two system changes have a marked effect on the privacy paradox?

Interestingly, this individuality does not extend to reported/desired settings as has been dealt with in detail earlier. For example, one would expect that varying cultural values are placed on the concept of privacy. However, when splitting the students according to their international status (home student and not) there is no statistically significant difference with each group desiring the same protection, again as has been demonstrated for other groups earlier in the chapter.

These questions would contribute towards a greater understanding of how the system itself can make it easier to handle privacy and, if it does, examine if there be a significant change in the way it is handled by participants.

5.4.2 Back to the paradox

The previous section identified that the likelihood of discosure for all pieces of information to friends was a modal response of 1 across the whole sample (a large majority of participants are happy disclosing anything to their friends circle). This corresponds with the overwhelming response in the reported settings question were the majority of participants believed that “Friends Only” was their privacy setting for most groups of information in their profile. Hence, if the granular perceptions of privacy are taken as intentions to disclose to certain parties and are matched with the reported settings then it is clear that participants have a clearly defined behavioural agenda. They wish only to disclose information to friends and believe (or report) that this is the case through their privacy settings.

However, the results from previous sections show that this is not the case with lax privacy settings for the sample as a whole. Furthermore, while participants stated a fairly definitive stance on the sharing of information only with friends for all data granules, this is not

reflected in the size of participant's friends lists (nor types of people participants reported as adding to those friends lists). Hence, there is a further paradox here with regard to the definition and treatment of social spheres within social networks. If this is the case then participants must actually be concerned with what they disclose within the network rather than believe their settings are adequate. Indeed, even with the strongest settings applied the varying social spheres and extent of their presence suggests that unintended parties will still be an issue.

In order to illustrate this phenomenon further, some individual cases shall again be examined for the traits listed in this section showing some specific examples of paradoxical behaviour. First, a single information granule shall be selected for study; the granule of date of birth due to its even spread across the sample and the studied social spheres illustrated previously. The most obvious form of the paradox shall be studied first; that is for a randomly selected participant who reported that they are not likely to share their data of birth with strangers, how true was this reported perception. Take, for example, participant ID#320, a privacy fundamentalist, who states that they would only share their date of birth with friends and are very unlikely to share with other social spheres. This is reflected in their perceived settings where they are of the belief that everything is locked to "friends only" for all of their pieces of information, including date of birth. However, the paradox is clearly evident as the profile is very open (a bracketed score of 4) and, indeed, the participant's date of birth is also available to anyone viewing the profile. By focussing on specific granules and more stages of thought were privacy is concerned the potential causes can be inferred with a greater degree of reliability (which shall be done later in this work). For example, where this participant is concerned, the perceptions of privacy are consistent until the actual behaviour is recorded; hence, the paradox is enacted at system level where this consistency is broken.

Taking another example with another information granule; participant ID#39, a self-reported concerned user, stated that they would only share their home town/location with their friends and very unlikely to share with colleagues or strangers. They stated that their settings reflected this and were set to friends only and yet their privacy settings are very open and the home town/location clearly visible when their profile is viewed. Again here the paradox is demonstrated in greater detail and with greater specificity than before (by focussing on an individual piece of information – both disclosure and protection of that disclosure). However,

intention scores across the majority of participants was evenly spread regardless of their level of concern where the majority stated they only wanted to disclose to their close friends.

A static measure of concern, therefore, does not seem to be a reliable measure of intention within the social network. Indeed, for the reasons outlined previously in this thesis, the context of the response to the concern measure could differ from the context of actual behaviour. This has also been noted in other work within the field (Johnson, Egelman et al. 2012). So, while a participant can have a general idea of what privacy means to them this is not indicative of how they *want* to behave in the network itself. However, there is a disconnect between this intended behaviour and actual behaviour within the network system. This would suggest that there is an influence at the point of interaction either by the UI or due to the role of human cognition. Consider the theory of hyperbolic discounting (Acquisti and Grossklags 2004) that suggests users trade long term risk for short term gain. This will influence end behaviour only as there is no obvious risk-reward concept in the measure of privacy concern but is present when choosing to disclose. However, what would happen if the UI informed the user of this phenomenon, would their behaviour change?

Recall also, literature has noted that the paradox could be caused by a lack of privacy salience within the network (Houghton and Joinson 2010); that is the information necessary to make informed privacy choices. That salience could be information that draws attention to the risk of disclosure or information that makes it easier to protect oneself thus ensuring the paradox does not occur due to a lack of skill as it has been suggested (Fang and LeFevre 2010). Note, this thesis is not suggesting that the UI is the one and only cause of the paradox but it does seek to understand what kind of role it could play in influencing and solving it.

The paradox then as it exists within this sample can be conceptualised according to the following figure:

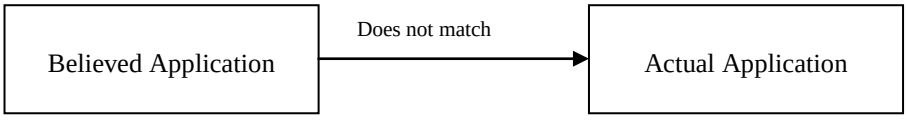


Figure 26 – Paradox overview

So what is concern? Within the context of this work it seems to be an arbitrary and general statement of worry about social network use. It does not relate to the context of actual behaviour; indeed, perhaps it cannot as a survey instrument cannot hope to model the

variables present in a decision to perform privacy behaviour. Hence, it could be that participants expect a level of engagement with the network and yet are worried about the effects they cannot control (trust in Facebook, invasions from nefarious parties and how the system will protect *them*). This may lead to the statement of concern seen in this study; it certainly seems that there is a baseline expectation to how one should act in a social network regardless of concern. It could also be that participants have stated a response that they feel is expected of them. They know they should be concerned, the media tells them so, and so they have responded as seen. All would point to a pre-arranged measurement of concern being ineffectual in predicting behaviour and that the paradox is caused by the contextual influence within which the interaction takes place (the system, user cognition etc.).

5.5 Justifying the Definition

Recall, from a previous chapter, Solove's definition of privacy is described as the driving definition behind this study. From this definition the focus was placed on the actual act of information revelation and, in social networks, this is particularly defined as the disclosure of personal information. Following the act of data revelation a variety of privacy related issues become possible making this particular definition an umbrella term for all those issues.

The survey contained one open-ended question dealing with defining privacy. This was designed to gauge what users of social networks define privacy as in order to create solutions to privacy challenges which are human centred. From the responses to this question, the following tag cloud was created of the words which appeared in the responses with the most frequency:

Figure 27 – Definition tag cloud

As can be seen from the picture the most prominent words used in the response to the question included: Keeping, personal, information and confidential which, fortunately, needs no work to turn into a workable privacy definition of "Keeping personal information confidential". Immediately this ties in with the first stage of Solove's privacy definition of information revelation and the fact that this is drawn from participant responses makes the definition more relevant to this study.

In order to reaffirm the point here; this definition would lend itself to one of limited disclosure. That is, certain pieces of information must be kept confidential and remain undisclosed in order for true maintenance of one's own privacy. Confidentiality is only truly possible if personal information is kept secret as the moment it is disclosed to a social network the boundaries the information resides within are potentially limitless.

This definition implies that there is an ownership of personal information that must be protected; the word "keeping" would suggest that information must be retained and not released in order to maintain that implied ownership. This in itself is paradoxical to the idea of social networks were disclosure is encouraged and once it has taken place ownership has shifted and is blurred. Interestingly, therefore, the definition is evidenced in the settings reported from participants in this study and the goal of social networks is clearly in evidence when looking at the actual, observed settings. Hence, this is a good summary of the paradox and the problem it poses; participants are aware of privacy and can form a clear, distinct definition of the concept as it means to them and yet act in an opposite manner to that definition.

5.6 Summary and Conclusions

This chapter has demonstrated several interesting points from the survey conducted and has done so with a greater degree of detail than previously available by taking both a case by case view and a view of the entire sample. However, the survey also provided empirical evidence for other areas of privacy which have long been stated within literature. The first of these is the idea of complexity within the privacy concept has been demonstrated; the variation between classifications of privacy concern has provided data for the changeable nature of privacy based on context. Furthermore, the related assumption of individuality was also demonstrated with empirical evidence through the granular privacy ratings and the differences shown between varying groups of people taking the survey.

From demonstrating these two points there is clear indication that there is no one-size fits all approach to providing technological privacy solutions; this is an assumption which has been made in literature (Rosenblum 2007). What is required, then, is a mechanism which facilitates this from an individual perspective, a mechanisms which enables participants to adhere to their own reported needs which, this survey demonstrates, are consistent across all concern groupings.

Furthermore, the generated privacy definition based on the open-ended question asked of participants has provided justification for focussing on privacy as the act of disclosure. This is useful for this study as the act of disclosure (and by extension protection of that disclosure) is clearly defined privacy related behaviour. It also raises further questions; first, why do participants take a clear view that sensitive information should not be disclosed yet do so anyway? Is this because they believe it to be well protected and available only to friends? Participants perceived settings would indicate this to be the case. Why, then, are these settings not accurately reported?

Interestingly the survey found that, regardless of the profiling used, the desired behaviour was consistent across all users (application of mainly "Friends Only" settings). Furthermore, in terms of perception and likelihood of disclosure there were consistent scores from across the sample that granules of information would only be disclosed to friends and not to strangers (with some variation of whether or not they should be disclosed to colleagues). This is despite participants freely admitting that they do not just add friends to their social networks. So what can be concluded from this? There is a clear need for granular controls when in a social network, participants only want their friends to know certain pieces of information and yet add more than just friends. Furthermore, their granular perception reflects this with varying ratings across the social spheres examined. Despite this, however, granular controls are not applied in the cases studied where the privacy paradox is clearly evident. Why this is the case is unclear and is the subject of further research throughout this work which attempts to examine the behaviour itself and why it is different to what users want at the point of interaction. The act of disclosure is arguably a more pressing issue for research based on these results as should controls be applied, participants still hold differing "real world" social spheres in the network.

It could be that users find it difficult to quantify potentially intangible risk and instead focus on the tangible outcome based reward; similar to hyperbolic discounting (see causes table in

chapter 2). For example, if the users has a goal of strengthening particular social ties with a specific friends group then the goal is clear and obvious (from a user's perspective); they know the message, intent and target audience. However, the potential risks are numerous and potentially unknowable leading to poor privacy behaviour that is in pursuit of reward rather than risk. This however, does not explain poor settings application as observed within this chapter which could be a result of low technical skill. It is the aim of this thesis to explore this point further.

Despite the range of perceptions of granular pieces of information participant consistently want, or report, friends only levels of privacy settings application. This could be due to it being the highest level of protection offered and therefore, the most reported as participant know that is what they perhaps *should* want their profiles to be given the media attention privacy and social networks receives. However, it is consistent with the samples perception of the friends only social sphere touched upon numerous times thus far; that is, the modal results for all granules is, indeed, very likely to share with only friends.

In relation to figure 10, the concern does not negatively correlate to intention or behaviour. However, this may not be paradoxical as the desire for strong privacy settings is consistent and perhaps should not be related to concern at all. The link between intention and behaviour, however, did increase with statistical significance and as such demonstrates the presence of paradoxical behaviour where it deviates from the desired or the intended.

To summarise, this chapter found the following;

- The privacy paradox is clearly still an issue and this survey has provided statistically significant evidence for it between participant's desired settings and their actual.
- Through examining the paradox granularly, greater detail is provided for its make-up where it is demonstrated that the disconnect lies specifically at a behavioural level; in particular, the importance of clearly defined social sphere has been demonstrated.
- It can be assumed that the same effect is occurring equally on all users given that their settings scores are not significantly different from each other when profiled (despite levels of concern differing between groups the settings scores do not).

- Similarly, the desire for privacy protection is consistent across all groups where particular social spheres are concerned (the sample as a whole only want to disclose to close friends).

- The complexity and individuality of privacy is empirically demonstrated showing the need for a flexible solution that enables users to suit their own unique privacy needs.

The questions still left to answer, however, include the following;

1. Why does the consistent link between intention and behaviour break down when actually in the environment being considered?

2. Does the environment, therefore, play an active role in breaking this link?

3. What will the study of behaviour, when it happens, reveal about the paradox?

These questions shall guide the rest of the study in attempting to produce a more holistic and thorough understanding of privacy behaviour in web-based services with particular focus on online social networks. The role of the environment is an important one to study as behaviour does not match intention at a system level. Hence, more work is required to ascertain why people are not acting the way they report. Given that behaviour is not statistically different across user groupings, it is a reasonable assumption to make that the cause of the paradox (or one of the causes) is acting on all participants equally. As privacy behaviour (like any behaviour) is a reaction to environmental stimulus, the User Interface is an appropriate area to examine as one of the potential causes.

The remainder of this thesis shall examine the role the interface *could* play in influencing users during interaction with a focus on disclosure behaviour as these seem to play a more complex role in poor privacy behaviour. That is, settings can be applied but privacy may still suffer due to the potential recipients of that information. Furthermore, there may be a tendency for a user to consider themselves "safe" as they perceived their settings to be enough (especially, when they have not been implemented as perceived).

Chapter 6 – Revisiting the Theory

Chapter 6 – Models of Cognition

6.1 Introduction

The field of HCI has a history of utilising psychological principles to design and implement experiments evaluating the use of user interfaces for the performance in achieving the designed tasks. For example, applied psychology has been typically applied in order to understand the individual user's behaviour within a non-human environment and to supply performance models to the UI designer (Card, Moran et al. 1983). This study however, proposes using behavioural psychology to evaluate the cognition of behaviour with a view to understanding the privacy decision of users within a simulated social networking environment.

The aim is to examine the various behavioural models which have been created within their research field and implement the most appropriate for the context of this study. This has several benefits to the research as a whole; first, a well-tested and proven theory shall allow for clearer and well justified assumptions based on the acquired dataset. Secondly, not only will the relation of UI elements to behaviour be examined but also, with the backing of a cognitive model, a deeper understanding of privacy oriented behaviour in general can be gathered. Finally, this research proposes and tests a new method to evaluating the *effect* of user interfaces on human habitual behaviour using the field of applied behavioural psychology. Social networks have been identified as a persuasive technology and the area of PT proposes the use of psychology within it (Fogg 1998). As such a novel research method is proposed where the elements of behavioural models are used to guide experiment design and used in analysis to aid understanding.

This chapter shall therefore review behavioural models which have been typically applied to computer systems research and select the most appropriate with sound justifications; in order to achieve this, the causes of the paradox, as reviewed earlier, shall be taken into account as an appropriate model should relate to each of the causes found in some way. This is a vital pre-cursor to the design of experiments.

6.2 Behavioural Models

This section will introduce and examine the behavioural theories which may be applicable. Specifically, behavioural change theories are the focus of this section given their applicability in examining and explaining the reasons behind behaviour and proposing ways in which that

behaviour can be changed (CommGAP 2009). Furthermore, only the theories which are immediately relevant to the study shall be considered; relevance is measured by how easily the model fits the context of this study.

The first behavioural model to be considered is Social Cognitive Theory which focuses on the role of self-efficacy in the learning of skills through peer observations (Bandura 1986). Self-efficacy refers to an individual's confidence in their own ability to perform a particular behaviour which they have viewed others doing around them. Initially, this looks like an interesting fit to the problem area under study here; users have observed others within the social network disclosing information and applying settings and have followed suit. Also, such a model has been applied to the area of technology use where its applicability in the training of computer skills was explored (Compeau and Higgins 1995). However, while this does partially offer a background to exploring the peer effects on disclosure behaviour within social network the various other potential causes cannot be considered (e.g. the role of the UI; which this study is concerned with). Furthermore, in the study mentioned previously, it was found that, compared to competing theories, Social Cognitive theory performed less well in influencing performance and behaviour.

The Stages of Change Model is the next model to be considered as a potential theory to understanding the privacy paradox and privacy behaviour in general. This model describes behaviour as habitual requiring stages of intervention to change it (McConnaughy, Prochaska et al. 1983); for example, disclosure of highly sensitive information could be habitual within social networks (indeed, this would agree with the persuasive technology aspect of the area). However, the model is longitudinal in nature; that is, it requires constant observation within stages of intervention that take place over a longer period of time than is available to this study. Furthermore, it does not cover all causes which are listed in the literature review chapter and is a model more suited to clearly defined habitual behaviour such as addiction, making it unsuitable to application within HCI (West 2005).

This brings the chapter to the models developed by Ajzen starting with the Theory of Reasoned Action (TRA). The TRA (shown below) describes the factors which influence the intention to perform a certain behaviour and thus the actual behaviour it leads to (Ajzen and Fishbein 1980).

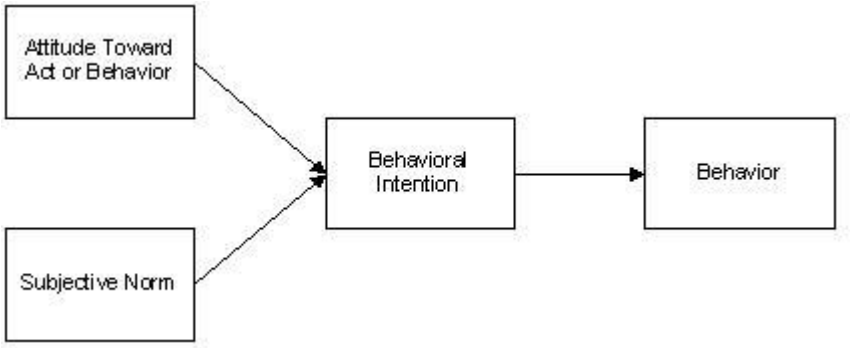


Figure 28 – The Theory of Reasoned Action (Ajzen and Fishbein 1980)

Attitude describes an individual's knowledge of the consequences of performing a certain behaviour which influences their intention to perform it. The other factor, subjective norms, describes the influence others play on the intention to act; for example, the effect of expert opinions or one's peers. Immediately the similarities to the paradox can easily be observed; attitude deals with user's skill or knowledge, while, the subjective norms category maps nicely to the effect of the user's friends within the network in propagating open privacy behaviour. Furthermore, Ajzen labelled these factors as salient features which are necessary in performing any behaviour and these are, usually, found in the environment the behaviour is being performed within. Recall, the lack of salient privacy information within the social network environment has been proposed as a potential cause of the paradox (Schneier 2009). However, a criticism of this model includes the fact that the intention to perform a certain behaviour is not always indicative of the actual behaviour which is acted (Sheppard, Hartwick et al. 1988). This criticism relates directly to this study where the paradox deals with precisely this kind of disconnect between intention and actual behaviour rendering this model inappropriate for use in the research of privacy behaviour.

In response to this criticism, Ajzen developed the Theory of Planned Behaviour (shown below) which added in the idea of control and perceived control (Ajzen 1991).

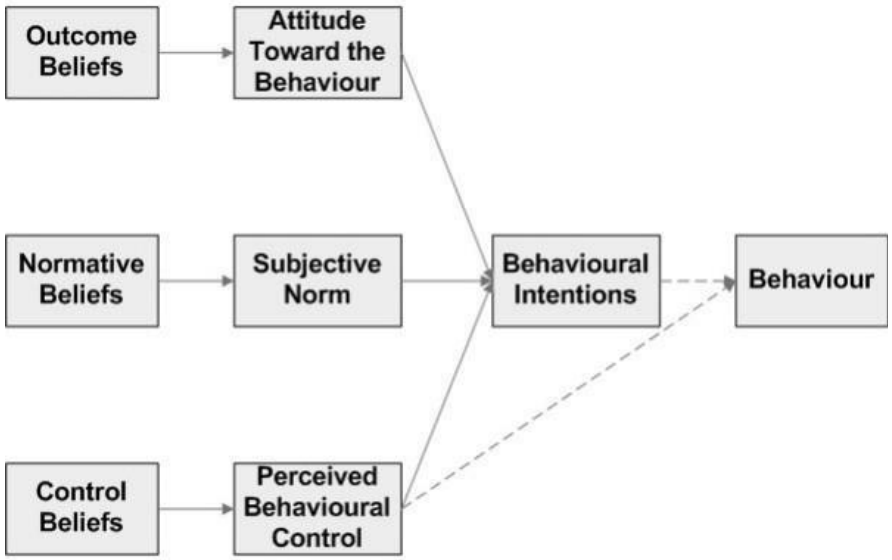


Figure 29 – The Theory of Planned Behaviour (Ajzen, 1991)

The added element of control here describes how easy an individual believes a particular behaviour is to perform and this perception influences the intention to act, while the accuracy of that perception influences the observable behaviour performed (represented by the dotted line). The addition of this feature makes this model much more applicable to this study within the context of the privacy paradox phenomenon. Added justification for the use of this model is shown with how closely the salient features proposed in this framework map to the potential causes illustrated in the literature review.

Specifically looking at each factor in isolation demonstrates how the model can be adapted to both explain and explore privacy behaviour; first, attitude (also called informed awareness) deals with the user's attitude toward privacy, their knowledge of privacy issues and their awareness of the context within which they are behaving (i.e. awareness of the consequences of disclosing a particular piece of information to a particular person, group of people at a particular point in time). Secondly, subjective norms (as with the TRA) deal with the influence of one's peers and experts where the behaviour is concerned. Finally, perceived control deals with the user's perception of how easy the network is to use in terms how easy it is to identify and protect their sensitive information. Interestingly, when this model has been applied to marketing campaigns it was found that a focus on only information dissemination was not as successful as campaigns which promote a positive attitude and ease of controls (toward the product in question) (Martiskainen 2007). Hence, improving only awareness of

privacy issues may not be enough to promote pro-privacy behaviour and why the UI could be so persuasive as it is closely related to perceived control which would seem to be the most influential factor.

The addition of control also relates to self-efficacy (closely tied to perceived control) which has been pin-pointed in other models as essential in performing a behaviour. In the case of social networks, a user's confidence in its use will be directly related to how they behave within it; so, the survey results seen previously could therefore, find an explanation through the theory of self-efficacy; namely, participants believe themselves to be confident, secure users of a social network which is not reflected in reality resulting in the paradox.

To offer a potential example; a user may believe to be aware of the consequences of poor privacy and report this accurately, however, protecting their privacy in the network (or identifying sensitive information) could be harder than they actually believe resulting in high levels of disclosure or poor privacy behaviour; the influence of perceived control would therefore be apparent. Another case may include a user who has limited knowledge of technology and privacy but knows they should be concerned about their safety due to increased media attention and panic; this may result in the reporting of high concern but little evidence of it in how they act.

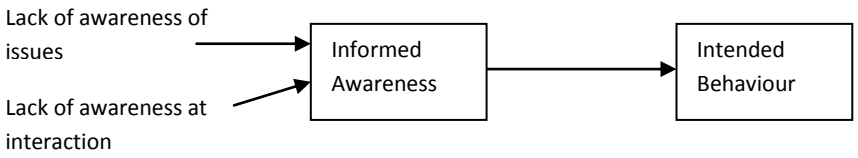
6.3.1 Salient Features

Each of these salient feature proposed by the TPB can be related to the causes stated in the literature review chapter (table 1). This section shall map these to the TPB to demonstrate applicability.

6.3.2 Personal Attitude

The personal attitude salient property described the influence that the individual's knowledge plays in the decision to perform a behaviour; specifically, knowledge and awareness of the consequences of the behaviour in question. Recall, the causes of the privacy paradox listed in the literature review as being personal to the user; users were said to be inexperienced of privacy issues (Kolter and Pernul 2009), unsure of the openness (Barnes 2006) or lacking in awareness of what is important to them were their privacy is concerned (Pötzsch 2009). Each of these issues or causes relating to the privacy paradox are covered within the umbrella feature of personal attitude as each of these are likely to inform the awareness of behavioural consequences. Furthermore, given the shifting and changing context of privacy within a social network (demonstrated by the granularity of information) it stands

to reason that this would also fall under this feature. Hence, the personal attitude aspect of the model can be translated thus in this privacy study:

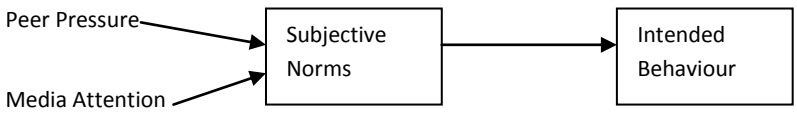


Privacy Personal Attitude

The dissection of this feature would suggest that if users of social networks are informed of privacy issues or reminded of their own privacy knowledge then there behavioural intention will change and they may enact better privacy behaviour.

6.3.3 Subjective Norms

The TPB suggests that behaviour is influenced by the views and actions of those around the actor. From the review of causes earlier recall that concern toward privacy issues could heightened due to intense media attention directed at the issue (Norberg, Horne et al. 2007) and users could act according to the behaviour of their peers (Strater and Richter 2007) (conforming to a kind of peer pressure: i.e. their friends are active so they must be as well). Again the fit with the TPB is well demonstrated as the subjective norms factor encompassed these causes well and provides a framework for them. It is proposed that the model can be expanded thusly:



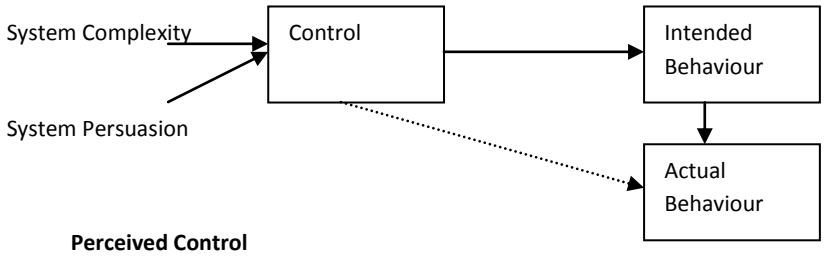
Subjective Norms

This suggests that behaviour is influenced by what user's friends are doing within the social network and their reported intentions may be influenced by media attention and the thoughts of experts in general. For example, a user may disclose much as their peers do so also and the theory of social capital would back this suggestion; as users disclose like their peers in order to strengthen social ties (Portes 1998). However, users could report being concerned about their privacy due to media and expert intention which would therefore give the reported levels of concern seen in privacy surveys (i.e. users know they *should* be concerned and state so).

The question which arises from this section of the model is: “will users act differently if they are aware of good privacy practices when they are acting?” as the TPB also specifies that behaviour is informed by expert opinion.

6.3.4 Perceived Control

The final salient factor is that of perceived control which deals with the perception of how easy the behaviour is to perform and how closely that perception matches reality. From the causes mentioned earlier it was suggested that social networks are designed to be open (Livingstone 2008), that they are persuasive in their design to encourage users to disclose information (Fogg 2009) and that the complexity of their design makes it difficult to control ones privacy (Heeger 1998, Brandimarte, Acquisti et al. 2012). This then is perhaps the most technologically influential factor; how easy does the social network make it to identify, control and protect sensitive information? It would seem then, that how easy users believe a technology is to use and how closely it is designed to promote privacy may play a role in influencing behaviour. Expanding upon and applying the idea of this factor to privacy gives the following mini-model:



The TPB therefore suggests that if users control over sensitive information is increased through identification and persuasion then they may enact better privacy behaviour.

Combining these into a complete figure gives the following:

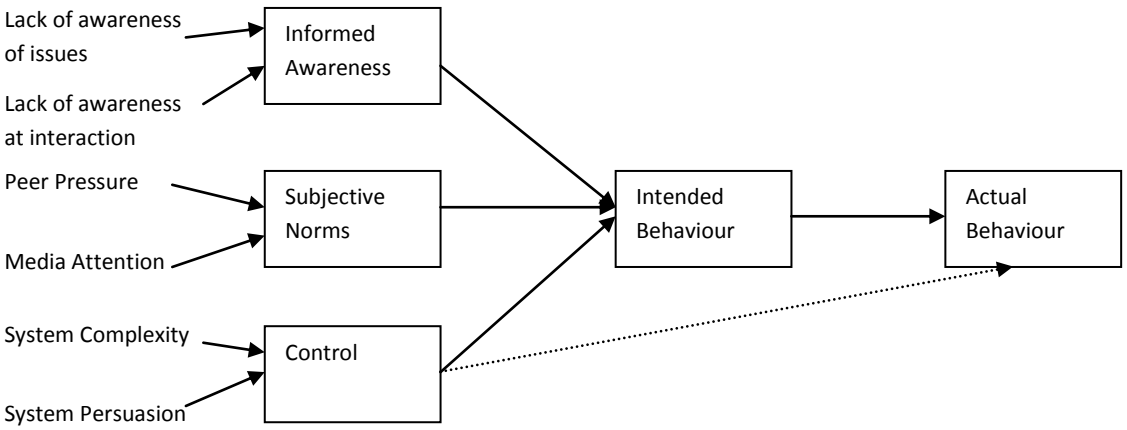


Figure 30 – The privacy paradox determinant factors

A note of self-efficacy: the TPB identifies self-efficacy as influential within the perceived control salient property. This deals with a person's ability to perform a particular task or behaviour (Bandura 1977) and is related directly to some of the causes listed in the literature review chapter; namely, awareness and level of skill within the network. There is work that suggests that users who exhibit tighter control over their information are more likely to seek out risk-coping information and mechanisms (Youn 2009). As such, user's self-efficacy may be improved through the TPB and then encouraged to ensure their privacy risk is minimised.

Such a framework as described in figure 30 can be used to embed the salient features within UI environments in order to examine if there is a relationship between the level of information disclosed and UI elements. This model offers several benefits to this research in such an exploration. First, the clear declaration of each salient feature allows experiments to be designed around each one offering a broader exploration of poor privacy behaviour. Secondly, a range of hypotheses can be used to pose questions of the research rather than just one adding a greater degree of validity to the data set and assumptions made based on it. Finally, the TPB has a solid and tested background in psychological experimentation as it robustly identifies the determinant factors of behaviour and shows how they relate to each other (Tonglet, Phillips et al. 2004).

6.4 Potential Criticisms

The Theory of Planned Behaviour is a model which has seen use in Information Systems Research before and has been criticised as being reductionist in its approach to understanding complex social issues. In response, this research is dealing with privacy *behaviour* which is a reduced element of the concept of privacy; any conclusions made will be based on the

behaviours observed within this study and only assumed to potentially be the case for the concept as a whole. Furthermore, privacy is acknowledged to be a complex problem and in order to study it, it needs to be simplified (Masiello 2009). This research is primarily concerned with behaviours relationship with the UI of social networks; therefore, an experimental approach within HCI is entirely appropriate and the TPB provides the ideal framework to this for the reasons outlined earlier (namely the clear description and bracketing of salient influence of behaviour).

Finally, in a review of competing models within Information Systems research showed that the Theory of Planned Behaviour provided the most specific information for developers (Mathieson 1991); another benefit for this work as specific, measurable information is required as HCI is specifically a development and testing field. The Technology Acceptance Model (TAM), for example, is a tool which primarily deals with self-reported perception of the ease of system use and how this correlates with the probability of system use (Legris, Ingham et al. 2003). As self-reported information is inaccurate (as shown by the paradox) and this model is longitudinal in nature (following a systems inception to its deployment and use), it is inappropriate for use in this study. Furthermore, the TPB was shown to provide a fuller understanding of behaviour itself by focussing on factors that influence it, due to its particular attention on control (Taylor and Todd 1995).

The TPB is therefore, highly relevant to this research as a framework which has the capacity to explain individual and isolated behaviour observed within social networks. When compared to other models, the TPB is specific in its characterisation of behaviour and the factors which influence making it ideal for application into experiments.

6.4 Summary

This chapter has chosen an appropriate framework for designing experiments which explore behaviour and its relationship to the User Interface within which it is performed. However, the model offered can also be used to provide an explanation to paradoxical phenomena observed, either in this study or in related works. It is a highly flexible mode due to its development within behavioural psychology dealing with behaviour in general and not specifically with technology use.

The salient features described by the model map nicely to the proposed causes found in related works demonstrating its ability to explain and inform work within the field of privacy

research and as a tool for designing experiments in HCI (dealing with any behaviour technology may produce). The following chapter designs experiments based on this model; through this a novel approach to embedding applied behavioural psychology into HCI experiments is demonstrated. By adding these salient features into the environment of the UI the ability of the system environment to persuade more ethical behaviour can be evaluated and any causal relationships between the UI (with salience embedded) can be tested.

Chapter 7 – Experiment Design

Chapter 7 – Experiment Design and Approach

7.1 Introduction

This chapter describes the design of a faux social network to be used in experiments with a focus on a sign-up process where participants are asked a series of questions in the creation of their profiles in such a social network. An outline of such experiments based on the TPB's salient features is proposed with an approach to implementing them in order to explore the research questions posed by this thesis. First, hypotheses are stated which this experimental procedure must explore.

7.2.1 Hypotheses

The model in figure 30 maps the potential causes found in the literature review of poor privacy behaviour and of the privacy paradox. Direct intervention at the point of interaction addressing these causes may promote more careful and considered privacy behaviour. Given that users appear to have a clear desire for privacy, it may be the case that reminding them of it will improve their behaviour. As privacy is observable through the amount of information disclosed and the protection applied (as outline in the literature review), it is reasonable to posit that these behaviours may be altered. Furthermore, given the definition from the survey chapter ("Keeping personal information confidential"), users may be influenced to keep more information private. As such, based on this model the following hypotheses are proposed based on Personal Attitude:

H1. A UI with privacy salient information aimed at informing or reminding participants of privacy consequences will influence user behaviour and decrease the amount of sensitive information they give.

H2. A UI with privacy salient information aimed at informing or reminding participants of privacy consequences will influence user behaviour and increase the level of settings applied.

Subjective Norms:

H3. A UI with privacy salient information aimed at informing participants of preferred privacy behaviour will influence user behaviour and decrease the amount of sensitive information they give.

H4. A UI with privacy salient information aimed at informing of preferred privacy behaviour will influence user behaviour and increase the level of settings applied.

Perceived Control:

H5. A UI with privacy salient information aimed at aiding users in identifying sensitive information will influence user behaviour and decrease the amount of sensitive information they give.

H6. A UI with privacy salient information aimed at aiding users in identifying sensitive information will influence user behaviour and increase the level of settings applied.

Again, in order to assess these hypotheses, measures of disclosure (what people say about themselves) and settings (how much they protect their digital identity) will be used as per the literature review.

7.3.1 Experiment Design

Participants will be asked to set-up their account profiles on a new social network built specifically for their University. Participants are recruited from the student body with the focus on Undergraduates as in the survey detailed in this thesis. During this process they will be asked a series of questions to "build" their profile and told that the more questions they answer the more accurate and complete their social network would be. These questions vary in *potential* sensitivity should they be answered by participants during the experiment. They also vary in the kind of response they require with text-boxes, drop down menus and radio buttons used to provide a variety of interactions. A full list of these questions can be found in appendix 5.

The set of questions are derived from the questions which were used in the survey from the first phase of this study. This maintains consistency throughout the study and adds to the generality of the results obtained since they will mirror real world systems. Furthermore, included within these questions, which the user may expect as part of their past experience, will be more sensitive questions covering a range of topics. The idea being to test the limits of disclosure based on the participants privacy perception of the information requested. The questions to be added are partly adapted from the Brandimarte *et al* (2012) study where a variety of questions were asked of participants with varying degrees of sensitivity (split into three categories) and differences between groups measured (groups had a variety of extra options for answering questions). For example, sensitive questions would be added dealing with drug use, drinking habits and professional conduct.

After answering these questions participants are asked to set their privacy settings. Furthermore, an extra link will be present on this page for the participant to follow and set

their separate connection settings; this is kept optional in order to explore how engaged with the system participants are and to see if they will seek out privacy protective mechanisms.

The experiment procedure will follow a between group, post-test analysis design consisting of a control group and three treatment groups based on the salient features described in the previous chapter and designed in the following section. The number of questions answered, which questions were answered and the settings applied will be compared between groups to examine if the treatments influenced participants to disclose less and protect more.

7.3.1 Control Group

The control group will mirror as closely as possible a real world application of a social network in order to make the experiments as ecologically valid as possible. Therefore, the first page of the social network sign up process uses CSS styling to match the feel and visuals of the first page of Facebook's home page screen. The experiment is created using HTML, PHP, CSS and JavaScript. The following two figures demonstrate the sign-up screen for Facebook and for the experiment:

Figure 31 – The Salford Network

And Facebook:

Figure 32 – Facebook

Notice the similarities between the two including the request for the same pieces of information to sign-up to the system. The mimicking of Facebook where possible shall increase the generality of the experiments as there is a real world example of the choices made and also address the ecological validity of the experiment where experimental settings should resemble real world systems so the design itself does not affect results (Lew, Nguyen et al. 2011). The initial experiment has been named the "The Salford Network" in keeping with the brief given to participants about a new network for students at the University of Salford; note, a second experiment is also conducted using a similar procedure as outlined here but aimed at students at Nottingham Trent University. This first page has a total of 7 questions asked of the user, two of which are optional. Completion of this initial page leads to the "profile builder" (figure 33) where the questions dealing with disclosure are found and completed by the participant; the idea being to answer the questions in order to create enough information to create a like network with their peers (that is, participants are informed that the more questions answered the more precise their resulting network with others will be).

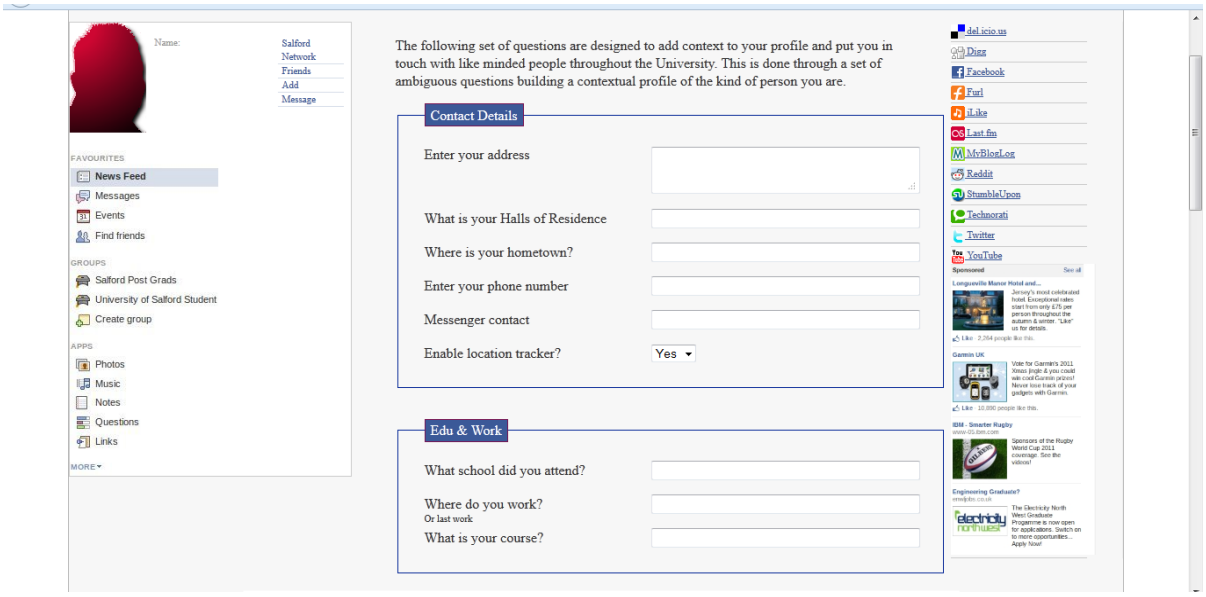


Figure 33 – Experiment Questions

The questions are split in to a variety of subsections with a specific focus; these are derived from Facebook where possible and some question groupings added for the experiment to test the limits of disclosure. For example, the highly sensitive questions form their own group which "add context" to a participant's profile (see following figure).

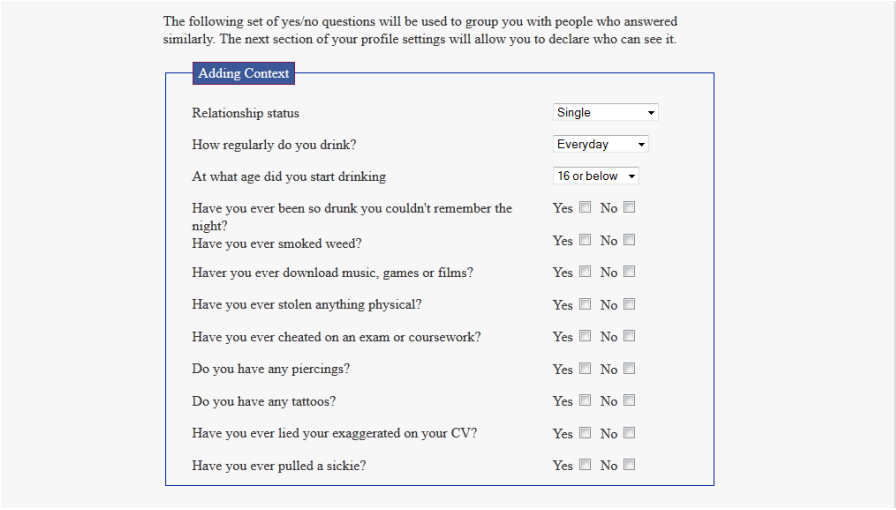


Figure 34 – Context questions

This page has a variety of interaction types including checkboxes, text fields and drop down selections. The idea being to test a range of interactions with a view to examining if the varying types influence the decision to disclose; for example, the above are most sensitive

but, perhaps, the easiest to answer so will this increase the likelihood that they shall be disclosed?

The final question grouping deals with marketing information and other pieces of data which may be of interest to advertisers:

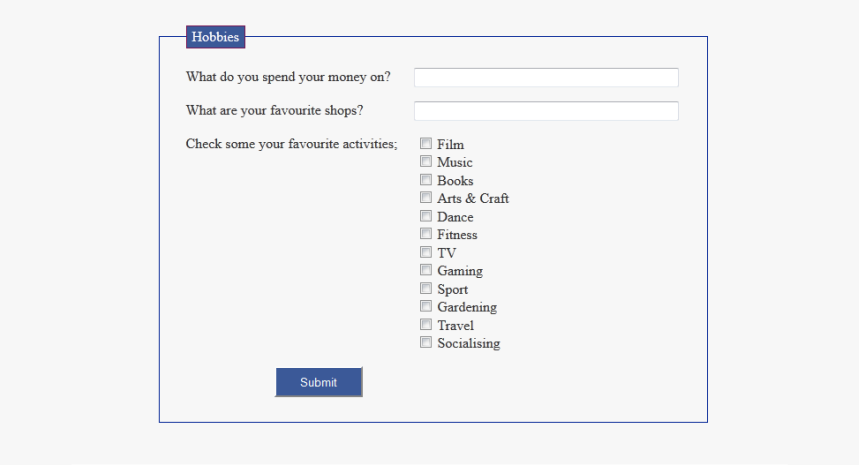


Figure 35 – Interests questions

Table 11 summarises the questions and their groupings with the experiment. The total number of questions answered by participants will be compared to the control to test H1, H3 and H5.

Table 11 – Experiment sections summary

Group	Detail
Sign-up	The first page of the experiment contains basic sign-up information and is not particularly sensitive.
Contact Information	The user is asked for contact information; such as address, phone, email etc.
Education & Interests	The user is asked for educational background information and personal interests.
Contextual Information	This section contained a variety of questions dealing with the personal context of the user. Questions ranging from relationship status to drinking habits were included in this section.
Marketing data	This section asked the user about shopping habits; favourite stores, items usually bought etc.

The next screen sets the privacy settings for the participants newly created profile. As mentioned previously, these are split into two groupings, one of which requires navigating through an extra link in order to access in order to examine participant engagement (figure 36). These settings mirror Facebook (both in content and separation) in order to promote familiarity. However, these do govern some information that was not asked of participants in the previous "profile builder" section; for example, wall posts and photos. This is because settings are often considered separately from disclosure and research suggests that they are not related (as may be assumed) (Christofides, Muise et al. 2009). The survey chapter also demonstrated that regardless of concern or perception the majority of participants favoured the application of the highest level of protection in their privacy settings.

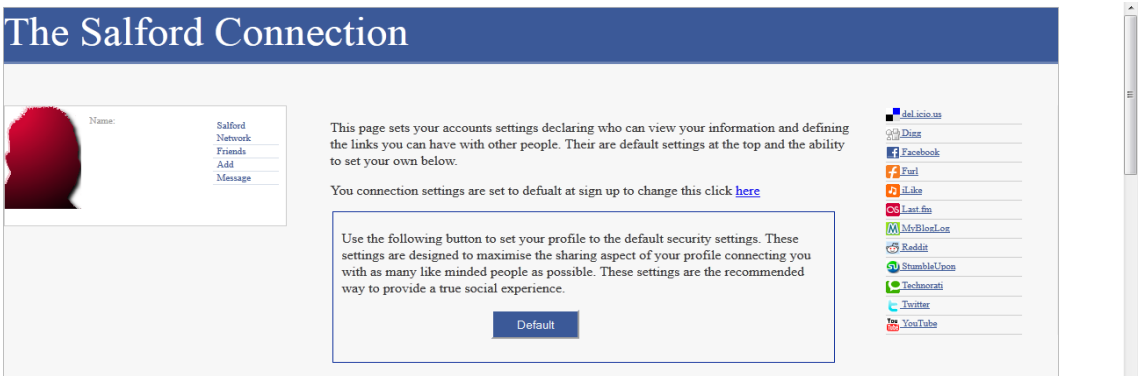


Figure 36 – Experiment settings direction

A default button is included should the participant wish to skip this process and apply open privacy settings. Figure 37 details the separate connection settings which deal with who can contact interact with the participant in terms of friend requests, likes and other contact information. Responses to these match those found in Facebook at the time of the experiment: everyone, networks, friend of a friend and friends only. Figure 38 deals with the profile information settings in terms of the personal information of the participant and is the dependant variable used to test H2, H4 and H6.

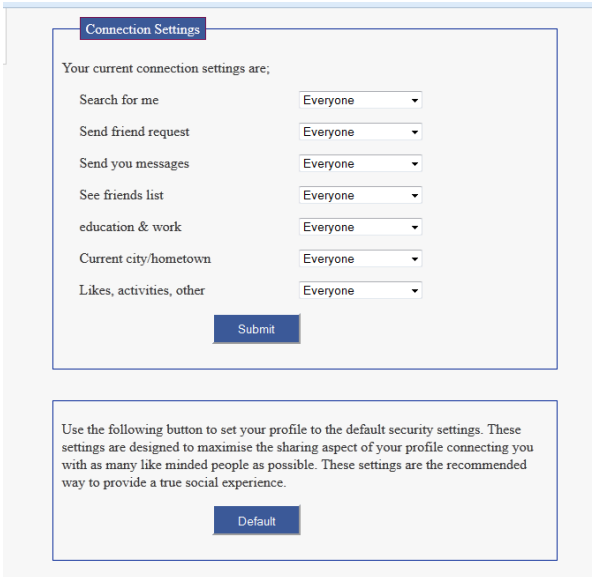


Figure 37 – Experiment connection settings

And the profile settings where the responses are everyone, friend of a friend and friends only:

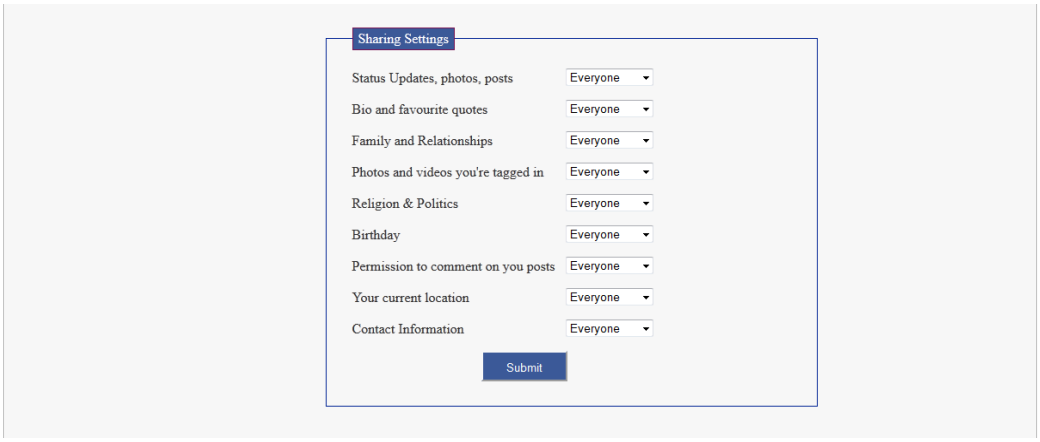


Figure 38 – Experiment privacy settings

The settings asked of the participants are summarised in table 12.

Table 12 – Settings summary

Settings	Detail
General Settings	These were explicitly presented to the user during the profile creation process and covered areas such as photos, education, interests etc.
Connection settings	These settings required extra exploration by the user (although the link was presented to them during the process). These covered the general settings of who can contact them and how visible they are.

Following completion of the settings applications the experiment in over and an exit survey is provided. Full details of experiment procedure are provided following the design of the treatments.

7.3.2 Personal Attitude

The Personal Attitude (PA) treatment derives a UI feature from the TPB aimed at informing participants of the *potential* behavioural consequences should they disclose a piece of information or choose an open setting and aims to test H1 and H2. According to the causes from Figure 30, participants will either need to be reminded of their own privacy attitude (Pötzsch 2009) or their awareness of privacy issues informed (Jones and Soltren 2005).

The work proposes that a simple UI metaphor can be utilised to quickly indicate the potential consequences for particular behaviours by grouping the requested information according to sensitivity. UI metaphors are particularly useful as they can aid users in quickly developing the mental models required to correctly use a computer system (Marcus 1998). Hence, a simple traffic light system is proposed called "Privacy Lights" to classify data according to its potential sensitivity.

The "Green" grouping deals with low impact data that, at its worst, could lead to social embarrassment (Strater and Lipford 2008) or annoyance from advertisers (Johnson 2010) if disclosed. For example, a user's hobby may not be sensitive but could result in embarrassment depending on their social group or favourite films may lead to annoying, targeted advertising if disclosed.

The “Yellow” grouping deals with more sensitive but not necessarily illegal data granules. For example, employers may examine social networking profiles prior to offering employment (Miller, Salmona et al. 2011). As such, some information could cost users a job, e.g. religious or political beliefs.

The “Red” grouping highlights data that could be in breach of the law, either by the user or toward them. For example, data in a SNS profile could lead to identity theft (Donath and Boyd 2004). Or information posted could be used in prosecution; for example, posting pictures of underage drinking or drug use (Morgan, Snelson et al. 2010).

It is important to note that these groupings are based on interpretation and the participant may not agree with them. However, the purpose is to suggest and inform participants and may remind them to enact their own needs or persuade them to follow the advice the treatment offers. Indeed, such a UI feature has been called a persuasive “suggestion” where an appropriate behaviour is mentioned at an opportune time (Fogg 2003). A full list of these groupings can be found in appendix 5.

Upon beginning the experiment the treatment is introduced to the participant in an informative pop-up (figure 39).

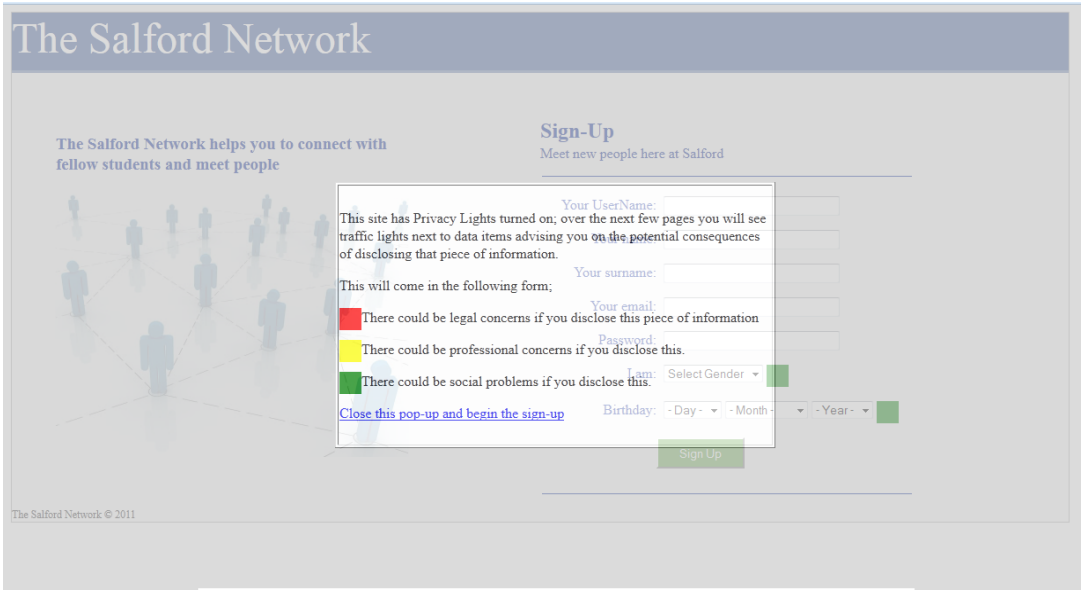


Figure 39 – Personal Attitude Intro

The lights themselves are present alongside each request for information as detailed in figure 40.

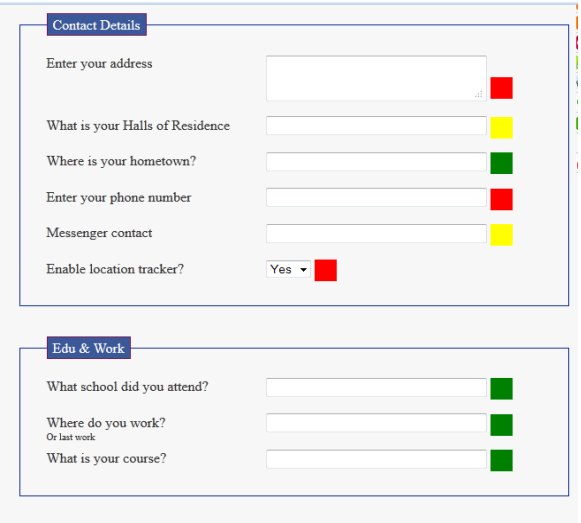


Figure 40 – Privacy traffic lights

Each of the questions asked of participants is given a "Light" which they may use to inform their behaviour. Each privacy setting is also given a "Light" based on the data the setting represents. When arriving at the settings portion the treatment is re-introduced with a settings specific pop-up box (figure 41).

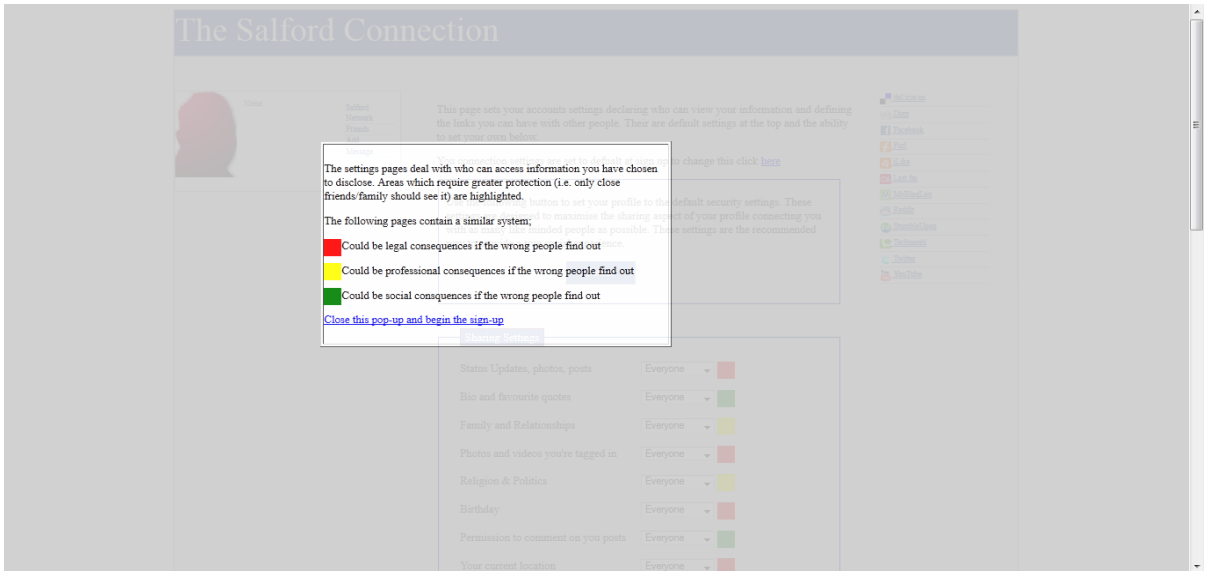


Figure 41 – Personal Attitude Settings Intro

Figure 42 demonstrate how the treatment is added to the UI for the profile settings.

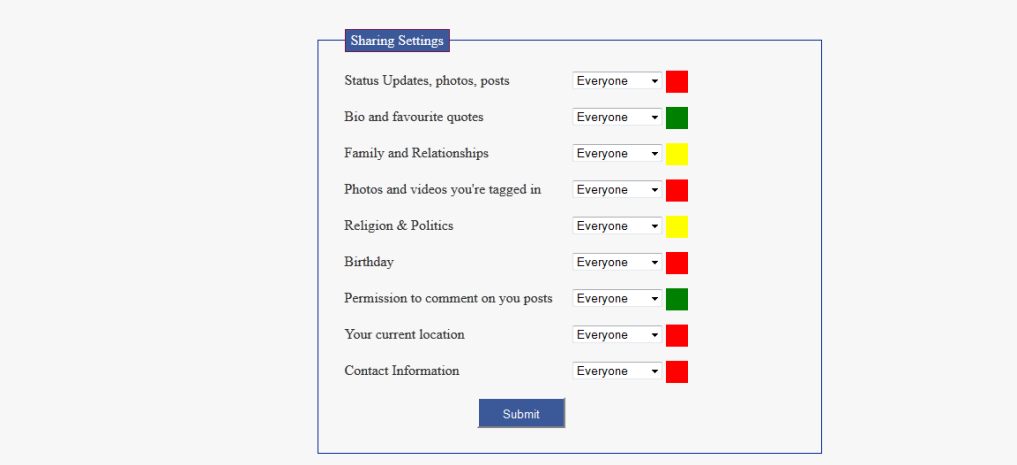


Figure 42 – Settings traffic lights

This treatment has added a simple UI metaphor based system to inform and remind users of the consequences of information disclosure and settings application. It is an easily recognisable colour scheme that participants should be able to quickly understand and assimilate. The treatment itself does not add complexity to the system itself so should not adversely affect participant efficacy toward the system.

7.3.3 Subjective Norms

The TPB suggests that behaviour is influenced by the thoughts or peers toward that behaviour by Subjective Norms (SN); for example, peer pressure and herding behaviour may influence users to act as their friends do. However, it also suggests that the opinions of experts also influence behaviour (e.g. from the media). Therefore, the treatment based on this property aims to introduce direct advice based on what others have done or suggest; one from other users and one from a "Privacy Expert" and aims to test H3 and H4. This will be introduced to the user within a pop-up box called "pAdvise" (figure 43) and "locks" the page requiring the participant to close the pop-up before continuing.

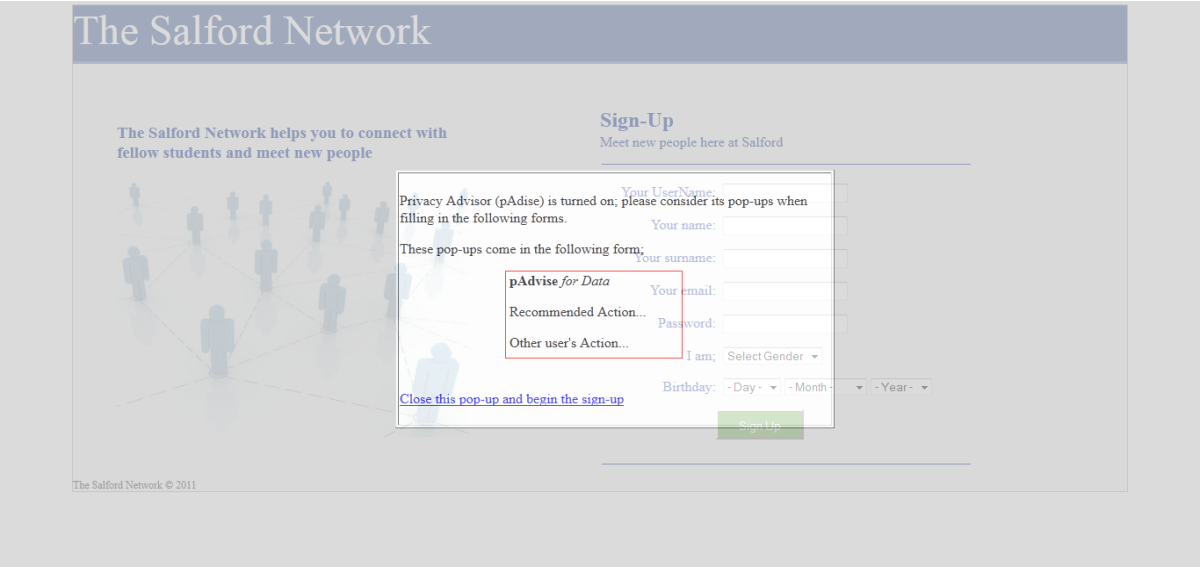


Figure 43 – Subjective Norms Intro

The peer advice states that other disclosed for all elements bar the "Red" category data granules defined in the PA group. The expert advice advises against disclosure for all bar the "Green" categories of information. A pop-up box for each question asked of participant is added by the treatment as illustrated in figure 44.

Figure 44 – Subjective Advice

The settings screen draws attention to the requirement for the separate settings in order to see if engagement can be encouraged (figure 45).

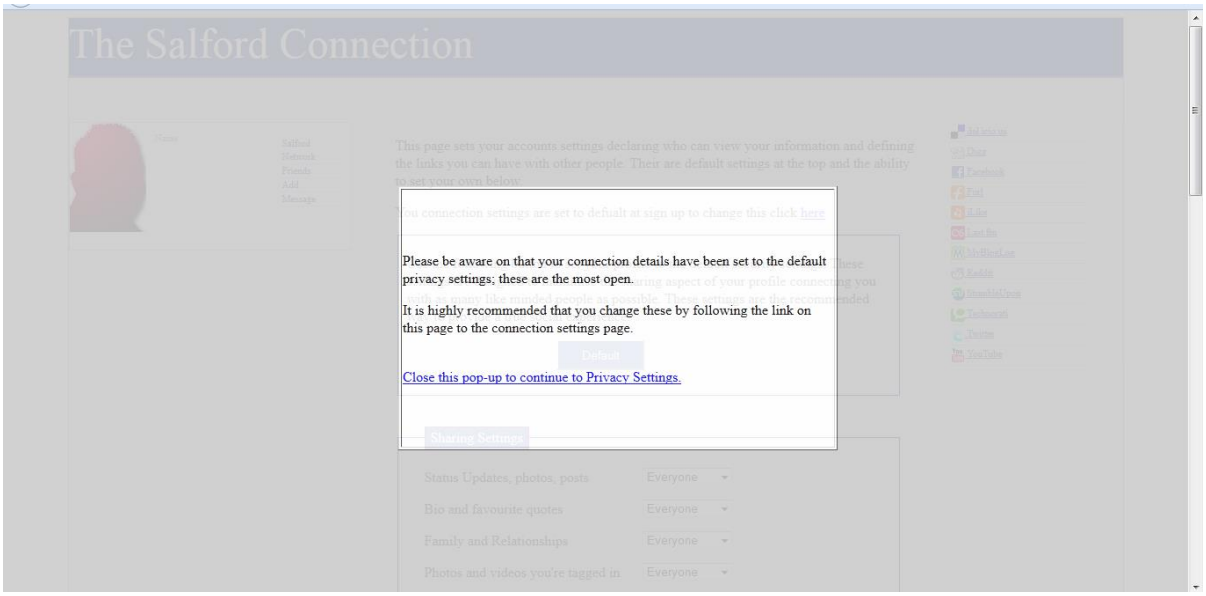


Figure 45 – Subjective norms pop-up

Finally, the settings themselves have similar pop-up boxes that advise on the recommended setting and the settings utilised by peers (figure 46).

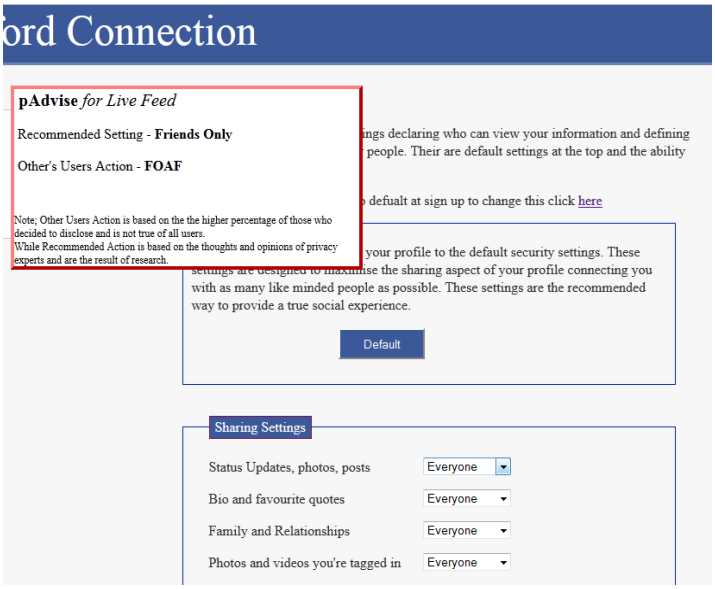


Figure 46 – Subjective Norms settings advice

This treatment has been designed to give clear and obvious advice through the use of a pop-up message box for each question and setting that the participant interacts with.

7.3.4 Perceived Control

The final treatment interprets the TPB's Perceived Control (PC) into a privacy salient UI feature and aims to test H5 and H6. PC suggests that the perception of how easy behaviour is to perform influences both the intention and action of behaviour. The causes listed in table 1

suggest that a participant's ability to identify and protect their information within a technological setting may produce poor privacy behaviour. Furthermore, research suggests that the concept of social networks themselves are designed to be open and users may associated them with the act disclosure (Fogg and Iizawa 2008). As such, the aim of the PC treatment is to aid the participant's control over their information by helping to review, identify and alter sensitive data. Therefore, it will give participants the opportunity to review and modify their data after each form submission by placing the data out of the social network context and into a privacy oriented one where the data with a (potentially) higher sensitivity is highlighted if the user has submitted it. Figure 47 illustrates such a screen which immediately follows the "profile builder" screen detailed in the control group.

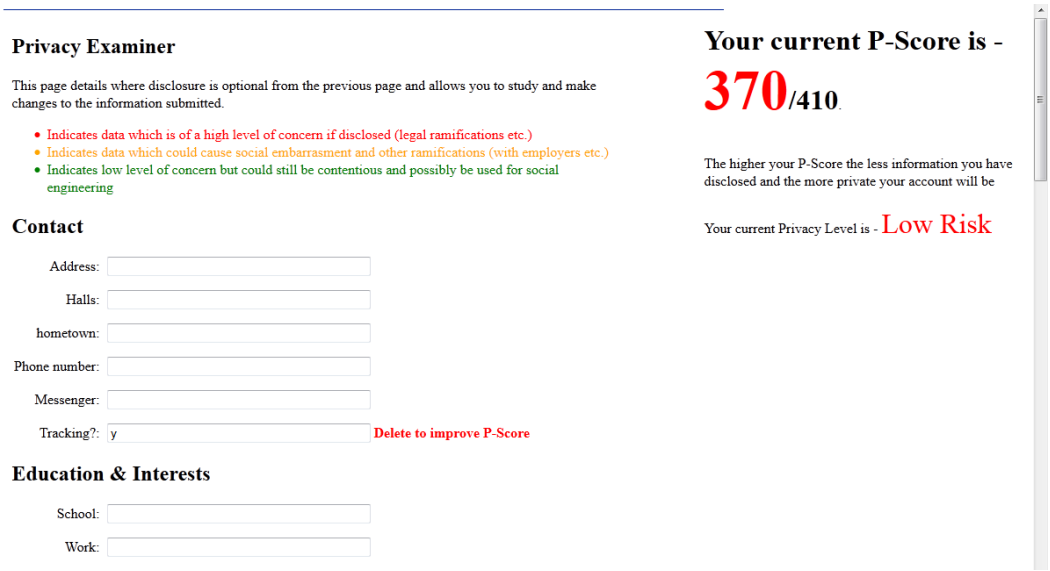


Figure 47 – Perceived Control Review Screen

To aid with the clear impact of interactions (effect of control) a dynamic P-Score is added and a level of risk given based on how much has been disclosed. The advice to delete in order to improve the P-Score disappears should a form element be empty and the P-Score reduces in increments based on the data deleted. The aim of which is to give a clear impact of interaction encouraging users to control their data with a tangible outcome. The forms prior to reviewing information (i.e. the social networking aspects) in this group is the same as the control group. To be clear, participants will access what looks like the control and, after each form is completed, they will review and modify their data in the treatment. There will therefore be two measures for this group: before review and after.

A similar screen follows the settings application which also compiles the settings into one screen. This is the only group where the separate "connection" settings will be explicitly displayed to the user during the compulsory flow of the sign-up process.

A full set of screenshots for the treatments can be found in appendix 4.

7.4.1 Experiment Summary and Procedure

This section summarises and outlines with clarity the specific procedure the experiment follows to allow for replication and verification. The above has outlined the groups to be included in the first experiment: a total of 4 groups in experiment one following a single factor, between groups, post-test analysis design. Using the TPB has the added advantage of defining multiple groups for the exploration of multiple potential causal links aiding the internal validity of the experiment (Church 2008).

Participants will be recruited from undergraduate students using a convenience sampling approach. It is noted that this may not produce a sample of a social networking population that is representative. However, this age group, as noted in the survey section, is the most susceptible to the privacy paradox and is therefore suitable in this study.

Participants are invited to volunteer to sign-up to a new social network and are approached at the end of teaching labs where they can take part in the experiment. Participants are given a simple set of instructions (which can be found in appendix 6) that directs them to the start of the experiment. At this stage, the true nature of the experiment is not revealed in order to minimise the impact of the Hawthorne Effect and to ensure the external validity of the experiment (Smith, Milberg et al. 1996).

Participants are randomly assigned to one of the groups in the experiment where they then follow the sign up process. Each group is asked the same set of questions and has the same base UI with the treatment groups SN and PA adding to that UI and PC treatment adding extra review pages. The experiment is stored on a web server and given a web address for access, participants therefore access the experiment through a standard web browser.

In the Control, PA and SN groups participants will complete a minimum of 3 pages in the experiment, possibly 4 if they follow the extra link to the connection settings. The PC group will complete the same number of pages as the control but with a review screen after each form submission as outlined in the treatment design. Therefore, there is a minimum of 6 pages the participants will complete, potentially 7 if the extra link is followed. After each form submission, the submitted data will be inputted into a MySQL database for analysis. If the

extra link is followed a record is also made in a MySQL database. As mentioned earlier, there are two measures for the PC group before and after reviewing their data allowing the exact effect of the review screen to be examined.

Measures are taken for what participants disclose (i.e. what questions they answer) and the level of settings they apply. In order to test H1, H3 and H5 the amount of disclosure across the groups is analysed for statistically significant difference when compared to the control. Should participants be considering their privacy it is assumed that the more sensitive questions will be answered the least. Therefore, data submitted by participants will also be split into the pre-defined categories of data sensitivity to examine difference. It should be noted here, that participants may not agree with the information groupings. However, they may still be persuaded by the treatments to follow the advice or improve their privacy score and will therefore choose to leave some answers blank.

In order to test H2, H4 and H6 the settings applied by participants are given a score based on their selection in the profile settings portions of the experiment. The profile settings have a score of 10 for FOAF's and a score of 20 for Friends Only with a total score of 200 possible based on the 10 settings that can be applied.

A similar score is generated for the connection settings with a score of 10 for networks, 20 for FOAF and 30 for Friends Only with a possible score of 210 for the 7 settings present; although, these are to explore engagement with the system and self-efficacy. Again, the scores in the treatment groups will be compared to the control for any potential statistical significance. This multi-measure approach (disclosure and settings) is intended to avoid mono-measure bias and improve the construct validity of the experiment (Oulasvirta 2008).

The experiment should take 20 minutes to complete and a short exit-survey follows which is aimed at examining if concern is notably influence by the privacy salient features participants are exposed to (designed later in this chapter). A number of participants will also be selected to take part in a small interview about their experience taking around a further 20 minutes.

7.5.1 Additional Approaches

The previous methodology sections outlined the need for a partial mixed-method approach to overcome the initial limitations of a purely quantitative one; furthermore, the need for an exit-survey was also mentioned to tackle the element of intention. As such this brief section shall outline the extra approaches implemented along with the above experiments.

Following the experiments is an informal interview with some of the participants. The goal again is to add a greater degree of richness to the quantitative data gathered with some of the qualitative opinions of the participants in the study. This shall be informal in nature and does not necessarily have a set agenda for the interviews as unforeseen talking points could be present through the observations. However, a set of general question to guide the interviews will focus on the participant's perceptions and thoughts of the salient features added. For example, did they affect their perceptions and thoughts on social network privacy?

7.5.2 Exit Survey Design

Following the experiments is an exit survey with the aim of tackling intention as the experiments themselves can only measure the resulting behaviour. This section shall describe the design of this brief survey.

The survey itself shall be hosted online and use Google Docs which has the added benefit of being linkable straight from the experiment website itself. Given that the aim is to explore intention and that the salient features have been based on the TPB, the survey questions shall be based on past surveys which implement the TPB as its driving theory. First however, in order to maintain consistency with the early portions of this study, the Westin privacy survey shall be included in order to get an overall view of privacy perception as they have been measured thus far.

Questions then shall be added to measure intention, attitude, control and the subjective norms of the participants in the experiments. First, intention deals with their view of disclosure and how comfortable they are with the concept. As such, the following set of questions deal with measuring disclosure developed from surveys used in health research and follow guides published for just such a reason (Francis, Eccles et al. 2004);

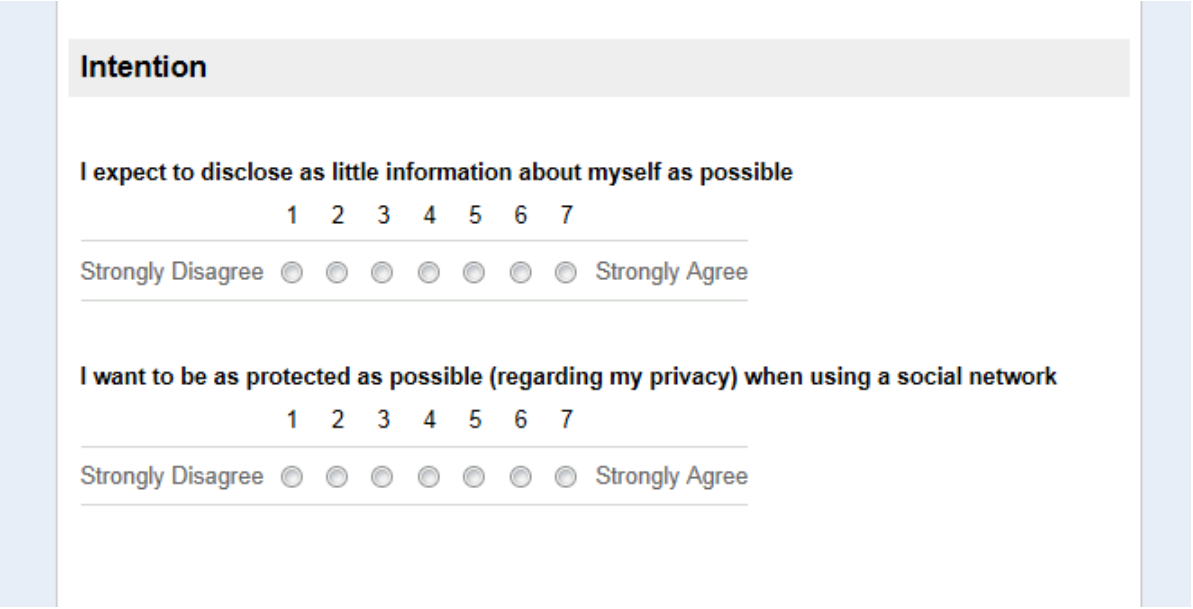


Figure 48 – Intention Measure

Typically, survey based questions which utilise the TPB take the form of a Likert scale response to a statement related to the perception being measured (Beck and Ajzen 1991). In this case, a general statement of intention of disclosure is made and participants respond with their measure of agreement. There are various methods for taking such a measurement within the scope of the TPB and the above is the most commonly used as it is the least resource intensive (Francis, Eccles et al. 2004). For example, a third method includes a simulation where a series of scenarios are used and perceived behavioural responses to them measured. Such an approach requires careful planning for reliability and is unnecessary within this study as the participants will have just been through an experimental scenario. A further alternative method focuses on performance; for example, asking how many patients out of ten would be sent for an X-Ray and comparing this to actual performance. While it may seem this is appropriate here, a more general question is required as the survey cannot deal with specific pieces of information (to avoid participant fatigue) or their varying contexts. As such, a question is required that deals with a participants perception of intention to disclose information about themselves in a more holistic way.

The idea being is to take a general view of the participant's perception of disclosure as described by the intention measure of the theory. Such an approach is taken throughout the design of the questions for each salient property being measured in this survey.

Measuring attitude involves using a set question and measuring the attitude using a varying range of responses. These responses are given a score and a final score calculated for a

measure of intention (the results added up and averaged). This has resulted in the following set of questions measuring intention:

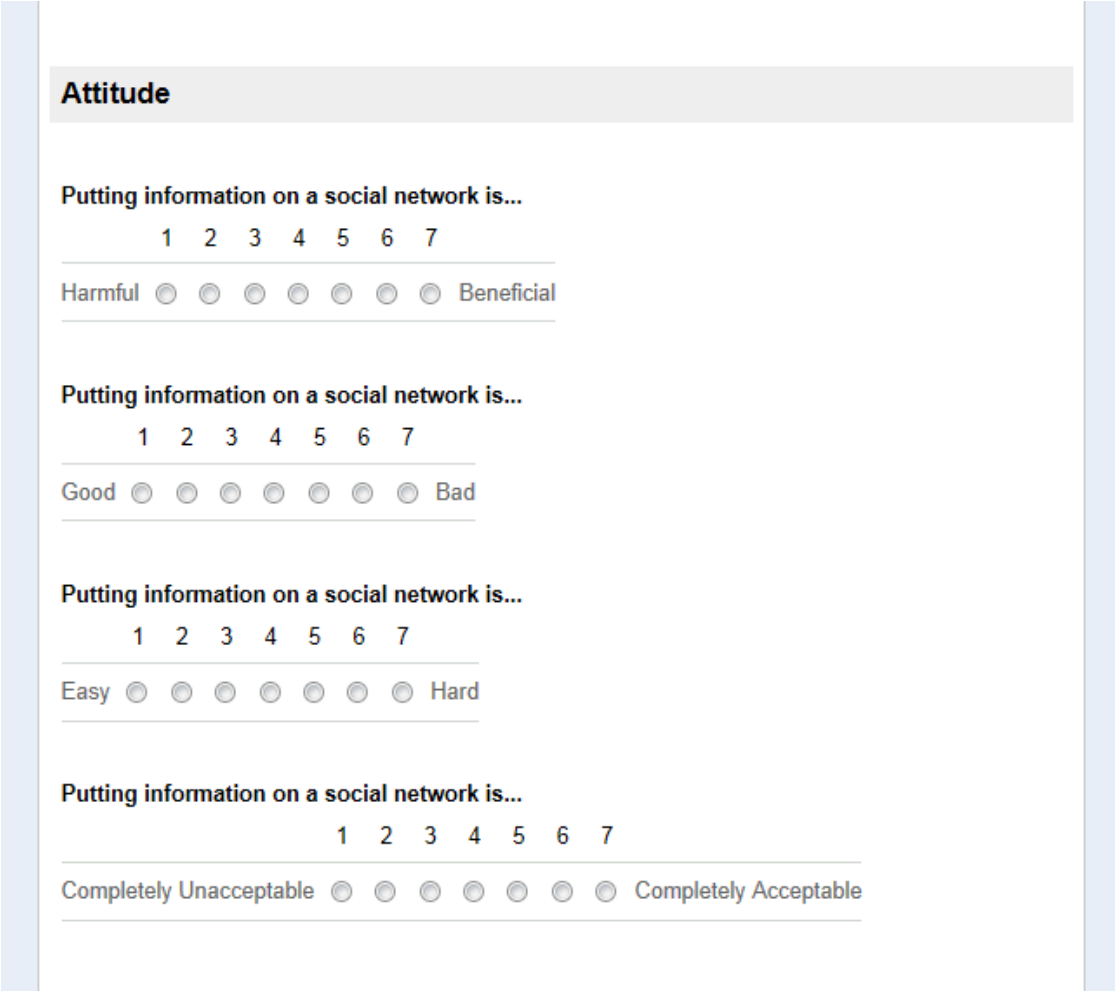


Figure 49 – Attitude Measure

Note, the switching of the negative end-points on the scales in order to avoid the risk of set responses or ticking through by the participants.

The subjective norms measure implements a similar question set looking at the propensity of participants to listen to the advice or actions of others. Based on this and on literature that have implemented such survey questions the following has been developed:

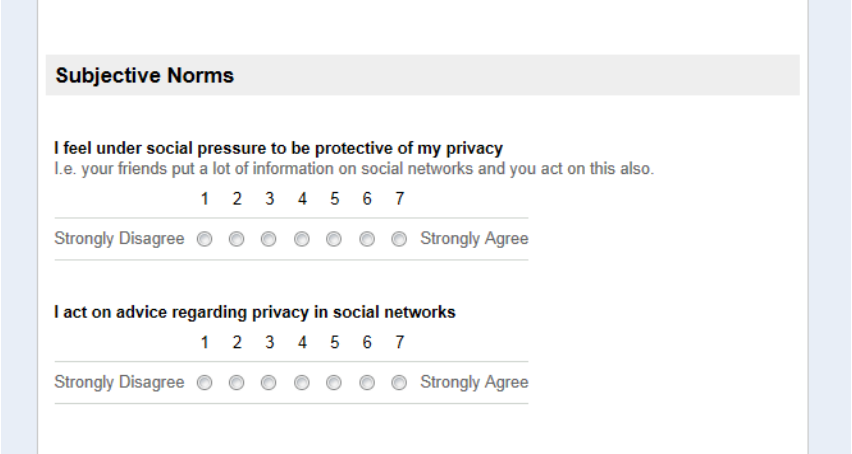


Figure 50 – Subjective Norms Measure

Again, the two are added and an average taken for a subjective norms score which is indicative of the participants likelihood of listening to others.

Finally, measuring of perceived control deals with the participant's perception of the data the experiment dealt with, whether they thought it was easy to control their information through identification and control. These concepts carry straight through to the following questions:

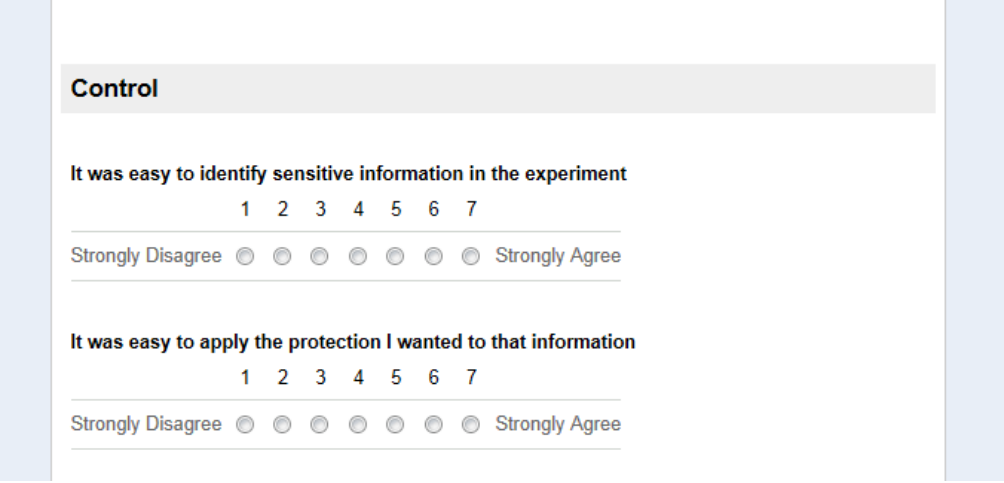


Figure 51 – Control Measure

The resulting scores are added and averages for a final score measuring perceived control where a higher score shows that participants thought that control over their information was easy.

This brief survey based much of its design on a manual developed for health researchers where the use of the Theory of Planned Behaviour is common (Francis, Eccles et al. 2004). Questions have obviously been adapted for use in this study. However one limitation should

be noted; that in usual implementations of this survey style and individual behaviour is the subject of each question rather than the general way in which it is used here. The adaption has been made in order to have a brief survey with the aim of getting a general measure for each factor in relation to the experiment and a longer survey would run the risk of participant fatigue given their involvement in the experiments prior to the survey. As such, the use of such a survey here can be considered a test of method and the appropriateness of its use shall be examined in the conclusions chapter.

7.6 Summary

This chapter has outlined a method of exploring privacy behaviour during a sign-up process to a faux social network. Participants will be split into four groups including three treatments and a control. The amount of questions answered each treatment group will be compared to the control to test H1, H3 and H5 while the settings applied at the end will be compared to the control to test H2, H4 and H6. Furthermore, a separate "connection settings" is available for participants to apply extra settings, the aim of which is to explore the engagement with the overall UI and the self-efficacy of the participant group.

Chapter 8 – Experiment One Results and Discussion

Chapter 8 – Experiment One

8.1 Introduction

The following chapter shall present and discuss the results from Experiment One designed in the previous chapter. Ultimately, the study is concerned with examining difference between treatment groups when compared to the control in terms of the levels of disclosure observed. This difference is hypothesised as being the levels of disclosure exhibited by participants in the experiment groups in terms of interactions with the system; i.e. how many questions did they answer. Recommendations are made for reworking the experiment based on the data gathered here and the subsequent methodological review.

8.2.1 Overview of Results

A total of 45 participants where gathered for this phase of the study from courses across the University of Salford. A full breakdown of participants within this experiment can be found in appendix 7. Specifically, these participants came from the Business School (E-business and Information Systems) and the School of Science and Technology (Media and Networking). Participants were randomised between the treatment groups.

Table 13, below, summarises the results of the levels of non-disclosure across all the treatment groups. Note, there are two measures for the perceived control group, the first summarises the initial behaviour before it is reviewed and resubmitted as a result of the treatment applied to the group.

Table 13 - Summary of Results

Group	Number of Participants	Average Total Amount of non-disclosure per participant	Standard Deviation	Total % of questions answered for the whole group	% of Yes answers when responded (total answered)
Control	10	3.7	4.3	87%	53.4%
PA	11	9.73	6.74	68%	36.7%
SN	12	10.17	7.32	66%	47.1%
PC1	12	12.17	9.6	59%	41.4%
PC2	12	18.58	7.54	38%	54.1%

Note, PC1 and PC2 represent perceived control prior and after privacy review and is the same group

The table presents the average number of unanswered questions per question grouping per participant and gives the average per participant for the process as a whole as well as the standard deviation for the groups. The total number of non-disclosed items for the entire

group is also provided. Again, from the previous chapter, the Perceived Control group holds two measures, one before and one after the privacy salient review added by the treatment.

Table 14 - Summary of Disclosure - Only "Yes" Responses

Group	Number of Participants	Average Total Amount of non-disclosure per participant	Standard Deviation	Total Group Disclosure when only counting "yes" responses
Control	10	8.8	3.3	68%
PA	11	15.8	5.9	45%
SN	12	14.6	6.3	48%
PC1	12	16.3	5.5	39%
PC2	12	20.5	7.1	28%

Table 15, below, details the results of a series of Mann Whitney U tests performed on the levels of non-disclosure for the treatment group compared to the control. The Mann Whitney test is chosen due to the small sample size and as result of the data not following a normal distribution. Hence, nonparametric tests are chosen as the statistical analysis tool for this experiment.

Table 15 - Summary of Statistical Tests ran on levels of disclosure

Group	Statistical Test applied	Disclosure P-Value for unanswered questions	Disclosure P-Value using only "Yes" responses to binary questions
Personal Attitude	Mann Whitney U	P=0.029	P=0.005
Perceived Control (Pre Salient Review)	Mann Whitney U	P=.003	P=0.002
Perceived Control (Post Salient review)	Mann Whitney U	P<0.0001	P<0.0001
Subjective Norms	Mann Whitney U	P=0.025	P=0.043

Given that a value of $p<0.05$ is considered required for statistically significant results, an initial look at the table would suggest that participants within treatment groups answered significantly less questions during their account creation. This does not take into account which questions were answered by the participants but illustrates that, where treatments are present, less questions were filled in for the group as a whole. The second P-Value calculation does take into account what responses were made to the questions. Specifically, those with a binary response and deals only with yes responses (i.e. responding "no" counts as non-

disclosure); findings, again, indicate that there statistically significant difference between the groups.

It would appear that H1, H3 and H5 have therefore tested true after an initial review of the data obtained from this experiment and this will be returned to throughout the discussion to ascertain the potential cause of the effect. A detailed breakdown of the results will be provided through the discussion following this section.

The above, initial, review of H1, H3 and H5 would suggest that changes to the UI have altered the behaviour of participants in the treatment groups. Table 16 breaks down the responses to questions according to the sensitivity areas described in the experiment design chapter and used as the basis for advice offered by the treatments in other groups. The table details responses for both the total questions answered in each category and total when discounting questions where "no" has been answered. As the green category does not contain any yes/no questions there is no data for this section and no column for it.

Table 16 - Location of Disclosure (Note, red indicates the least answered sensitivity grouping)

Group	% of "Green" questions answered	% of "Yellow" questions answered	% of "Red" questions answered	% of "Yellow" questions, only yes	% of "Red" questions, only yes
Control	82%	90.3%	90%	60%	63%
Attitude	60%	75%	67%	42%	32%
PC1	54%	64%	57%	40%	33%
PC2	39%	42%	33%	30%	18%
Subjective	63%	76%	58%	48%	33.3%

If participants are considering their privacy then it is a reasonable assumption to make that disclosure will be the least in the more sensitive categories. The size of this effect should be greater in groups with salient properties embedded as they may draw attention to the sensitivity of information, providing participants with a "nudge" to consider their privacy. It would appear, from table 17, that participants did not choose to disclose based on their privacy due to the decreases in the green category also.

Table 17 - Statistical comparisons of categories to control

Group	Test	Green	Yellow	Red	Yellow, only yes	Red, only yes
PA	Mann Whitney	.082	.084	.022	.033	.001
SN	Mann Whitney	.131	.199	.004	.193	.005
PC1	Mann Whitney	.026	.003	.005	.014	.002
PC2	Mann Whitney	.006	.001	.001	.006	.001

When examining only "Yes" responses disclosure was the least in the yellow and red categories when compared to the control. However, the spread when examining total questions answered is less clear; for example, the PA and SN group is not statistically different from the control for the yellow categories while the PC group held a significant reduction in the green category also. This may suggest that participants are not disclosing based on the sensitivity of the information asked for.

H2, H4 and H6 posited that participants within groups with salient features embedded in the UI would exhibit increased application of privacy settings when compared to the control. Table 18 details the averages for the settings scores obtained across the groups.

Table 18 - Settings Results Overview

Group	Privacy Settings (St. Dev)	Connection Settings
Control	108	0
PA	145.45	0
SN	72.5	10
PC1	160	13.3
PC2	171.67	92.5

Table 19 details the results of Mann Whitney tests performed on the settings scores.

Table 19 - Settings Statistical Tests

Group	Statistical Test	Settings P Values
Personal Attitude	Mann Whitney U	P=0.468
Perceived Control (Pre Salient Review)	Mann Whitney U	P=0.381
Perceived Control (Post Salient review)	Mann Whitney U	P=0.228
Subjective Norms	Mann Whitney U	P=0.381

The treatments representing personal attitude and perceived control held increases in the average settings scores for their groups as a whole. However, these were not statistically significant as detailed in table 19. The subjective norms treatment actually held a decrease (trending in the direction opposite to the expected) in the group average when compared to the control and this, also, was without statistical significance. The reasons for this will be explored in the following discussion.

No groups followed the connection settings with any consistency suggesting that perhaps self-efficacy was not affected to the extent that participants were willing to explore further privacy enhancing mechanisms. This will be explored later in the chapter.

Therefore, the null hypothesis for H2, H4 and H6 cannot be rejected from this dataset as there was no evidence that participants protected more when reminded of privacy when compared to the control group.

8.2.1 Discussion & Limitations

The following discussion shall examine each group in turn by detailing the specific responses to questions; where appropriate information from observation and post experiment interviews shall be used to expand on discussion points.

8.2.2 Control Group

Participants within the control answered more questions during the experiment than the treatment groups; specifically, participants answered 87% of the questions asked of them within this group. If only including answers where the participants responded in the positive the total percentage of disclosure for the group stands at 68%. Interestingly, the standard deviations for the control groups' responses were the least when compared to the treatment groups (table 13). This would suggest a more consistent exhibition of behaviour from the participants within this group.

The questions asked during the experiment varied in sensitivity and the treatments group these into categories (described in the experiment design chapter). The treatments were designed to classify the questions and inform participants of them; the control group, however, were unaware of this overt classification and relied on their own perception of data sensitivity to fill in the questions. The spread of question responses across pre-defined sensitivities can be seen in table 16 which shows that participants actually disclosed less in the green category when compared to the yellow and red (if considering all questions answered). If only considering "Yes" responses, participants disclosed the least in the Yellow (marginally) category of data items.

However, answering a question in the negative can still be considered disclosure of the self in some way. If privacy is considered the right to form an identity to particular social spheres then answering in the negative will allow others to form an opinion of the individual (although, practically it may not be clearly privacy *invasive*). Indeed, the Theory of Social

Capital would suggest that anything disclosed about the self can be used to form an opinion about one's identity. For example, admitting to not drinking may put an individual at a social disadvantage when confined to a party oriented social group. However, participants may mislead and lie about their submissions and it is unclear if this may be a tactic employed by participants in this experiment.

These initial results may suggest that the majority of participants within this group exercised high levels of disclosure during their account creation and disclosed potentially sensitive information in the process. This high level of disclosure could be due to participants building social capital with their intended peers (Portes, 1998) or due to them being driven by the perceived goal of the system (to create their accounts) and therefore prioritising this over other considerations such as privacy which may suffer from the secondary goal problem (Bonneau, Anderson et al. 2009). There is also precedent within research to suggest that participants simply answer questions asked of them with little thought within social networking sites (Strater and Lipford 2008).

Observations certainly seem to suggest that participants are focussed on the questions at hand as a number of them seemed to scan each individual question with the mouse cursor before answering. Attention was focussed on the questions specifically rather than the system as a whole as exhibited by no participants changing any of the connection settings which required exploring an optional link. Indeed, experiments with eye-tracking software have demonstrated that participants tend to focus more on the point of interaction (Whalen and Inkpen, 2005). So, participants seemed to focus quickly on the questions being asked of them. Again, these questions can be considered the sub-goals of the primary goal of account creation and hence HCI would suggest they are ignoring potential distractions in pursuit of their perceived primary goal (Bishop, 2007). Furthermore, research has suggested that users tend to forsake consequences of what is being asked of them, ignoring all else (Wu, Miller et al. 2006). Such an observation is in line with the Milgram effect which describes how people tend to act according to the whim of a perceived authority figure despite their usual inclinations (Milgram and Fleisner, 1974).

When asked post-experiment why they interacted with the system in such a way participants responded: "I tried to give the system what it wanted in case it would not let me to the next section" and when why they answered questions: "because the system asked for it". This highlights a potential limitation in the experiment design as it may not have been clear to

participants that disclosure was optional; i.e. they had to answer the questions in order to finish the process.

Furthermore, interaction with the green category (which is not particularly privacy invasive) was less than the yellow and red sections (if including no responses). Work in security suggests that users tend to implement security features based on the ease of interaction (Sasse, Brostoff et al, 2001). Hence, this section may have required more user effort to fill in than questions with simple binary responses. Again, this highlights a potential limitation in the design as there are more of these open questions in the "Green" category than in the more sensitive sections. These points will be picked up throughout this discussion.

Finally, the settings obtained for the group held and average of 108 for the group with a standard deviation of 99.9. The highest level of protection was chosen by 50% of participants in the group (selecting "Friends Only" for each information grouping) with 40% of participants selecting no settings at all in this section of the process. So, when interacted with, participants tend to protect liberally as suggested as desired behaviour within the survey.

8.2.3 Personal Attitude Group

From table 13, the Personal Attitude group held a decrease in disclosure with statistical significance; both in terms of the total amount of questions answered and when only responding in the positive to binary questions (68% and 45% respectively). Interestingly, the lowest answered category was the less sensitive green grouping of questions; although the yellow and red also decreased compared to the control (with statistical significance for the red category). There is a larger difference when examining only the "Yes" answers with the red category of questions having the largest decrease in disclosure when compared to the control and the yellow in the middle as would be expected.

Interestingly, of those who answered the binary questions in this group only 36.7% answered in the positive. This is the lowest of all the treatment groups and could suggest that participants are utilizing deception to manage their privacy as a result of the salient treatment. However, it could also be due to variation within the group and it is unclear from this experiment alone if this is the case.

So, does a decrease in the green category (when compared to the higher sensitivity groupings) mean that participants are not being selective in what they disclose and simply not filling in the questions that take more effort? It is important to remember that the traffic lights

aim is to inform *or remind* participants of their own privacy desires. Hence, it could be that participants are unwilling to disclose data in the green category as they are private to them. For example, participant's favourite films could not match their peer group's expectations leading them to leave it blank. Indeed, the literature review chapter noted the problem of "new data" is social networks causing privacy problems for users that had not previously. While disclosing ones telephone number is potentially invasive in the traditional sense that information will not be used to form social opinions of the discloser but social information (films, books, etc.) could be (depending on the individual).

However, it could be that the green questions took too much effort to answer as they held more text boxes for answering in lists or questions that people may not have an answer to. Indeed, participants mentioned post-experiment: "(the lights) showed that it is ok to leave some questions blank". The suggestion being that, when the option to answer questions was clear, participants would then choose what to answer. It is, however, unclear if that choice is based on increase consideration of participant's privacy or for other reasons; although, disclosure was the least in the red category when considering only the positive responses.

In the post-experiment interview when discussing the contextual questions, one participant, talking about their CV, stated: "everyone is going to tick yes so I shall as well". Justifying their actions based on the behaviour of their peers in a phenomenon known as herding behaviour which has also been used as a potential reason for poor privacy behaviour (Gross and Acquisti 2005).

As in the Control group, no participant changed the connection settings. This brings into question the extent to which participants engage with the entirety of the User Interface and suggests that participants go straight to the obvious interactions of the form elements. Again, eye-tracking software may be useful in examining what participants look at in the UI and for how long. The main page of settings increased from the control without statistical significance. Of the participants who applied settings 55% applied friends only to all categories with the remainder being somewhat more selective but it would appear not in relation to sensitivity ratings of the settings. This would suggest a more consistent behaviour from participants within this group when applying their privacy settings. The relationship between disclosure and privacy settings is not clear from this experiment. Participants in this group, on average, applied a higher level protection than the control group yet disclosed less. Indeed, Christofides (2009) suggests that the two are different processes affected by different

aspect of personality and are not negatively correlated. Further qualitative work is required to explore this further from a participant's point of view within this experiment.

8.2.4 Subjective Norms Group

Levels of disclosure within the subjective norms group decreased with statistical significance compared to the control with participants answering 66% of the questions asked of them for the group as a whole; this is a similar level to the PA group and there is no statistically significant difference (p=.976 and p=.566). When only considering "Yes" answers the group answered 48% of the questions asked (a decrease of 20% from the control). Of the binary questions answered, 47% answered in the positive; this is only a 6.3% decrease from the control showing similar behaviour across groups. That is, there is a similar ratio of "Yes" responses which would suggest that participants are choosing to leave out question responses when answering in the negative.

The levels of disclosure decreased from the control across all sensitivity groupings with the yellow and green category of data being the most answered (Table 16). However, when only examining the "Yes" answers, the red category held the least amount of disclosure. The yellow category was the most answered with red being the least. This group did have the highest rate of disclosure in the green category (for the treatment groups). This is the only category where the advice from both the expert and others users matched providing some evidence that the treatment aided participants in disclosing in the least sensitive categories.

In the other sensitivity categories, the treatment tended to hold conflicting information, the aim of which was to ascertain which participants favoured. However, it could have had the effect of confusing users and introducing "signal noise" into the UI that prevents users from implementing their choices effectively. Furthermore, it is unclear from this experiment what the exact effect of the treatment is due to the conflicting information and the small sample size. Future iterations of the experiment should only include one form of advice to ascertain if it plays a role in persuading users to follow that advice.

Again, the ease of interaction played a role in deciding what to disclose for participants. One participant remarked in the post-interview: *it is easy to answer some of the questions but a pain to answer the ones regarding interest*; the other questions referred to were the very invasive questions. The participant felt that due to how these questions were asked, they were easy to answer (simply clicking yes/no with limited data entry); however, the interest

questions required time consuming attention regardless of the invasiveness of the information. The participant also stated that: *I know I shouldn't have answered those questions now but they were easy to fill in at the time*.

Finally, another participant stated: *I found answering difficult to approach, until I thought of it in the same way as Facebook*. This would indicate that there is an element of conceptualisation required when the system is new to the participant; that until they are comfortable with it they will do as it asks unless they treat it in the same way as a system they are experienced in.

The privacy settings for this group where actually lower than the control group, the only treatment group to demonstrate such behaviour and it is unclear from this experiment why this might be the case. This could be due to the extra information provided by the treatment increasing the UI signal noise and causing confusion from users. This would decrease their self-efficacy (confidence in behaviour) so a number of participant would leave settings at the default for fear of incorrectly administering behaviour. Similarly, the control paradox (Brandimarte, 2012) described how increased control increased disclosure perhaps due to the participants decrease in self-efficacy.

How may signal noise have been increased? Participant's perception regarding pop-ups seemed to play an issue. The initial message presented to the user (which locked the screen) seemed to cause some participants to believe an error had occurred. Upon realising the screen could be closed, participants did so quickly in order to continue with the sign-up; this then became an issue for all further pop-ups (they were not an error so could be ignored, thus avoiding further complexity) including the pop-up which attempted to draw attention to the section on the connection settings. Indeed, it has been noted that error messages within web browsers and web development in general often confuse users (Lazar and Huang 2003) hence if they have been interpreted as such then signal noise and confusion could be increased.

However, it could also be due to simple variation within the group due to the small sample size. A reiteration of the treatment is required in order to provide only one piece of advice allowing for comparisons to be made to this group.

The connection settings where only set by one participant in this group and where roundly ignored by participants as in other groups again bringing into question the extent to which users engage with a system when first introduced to it. It may require a degree of confidence

in using the system before it is fully explored and to gain that confidence prolonged system use is required.

8.2.5 Perceived Control Group

Both before and after the review of data performed by participants, this group held a decrease in levels of disclosure that is statistically significant; this group also held the largest decrease when compared to the control. Before salient review, 59% of the questions asked of the group where answered in total and this dropped to only 38% after participants data had been examined and manipulated. Interestingly, the ratio of "Yes" responses where the questions were answered increased 41.1% to 54.1% after salient review as a greater number of participants who answered "No" deleted their responses altogether. So again, the preference toward deception or withholding information is not clear from these experiments alone.

Disclosure decreased across all the groupings of information. First, looking at total question responses prior to salient review (PC1) shows that the green category was the least answered with the yellow and the red (table 16) being similar. After salient review, the spread across the groupings is fairly equal when considering all question responses which would suggest that participants are not considering their privacy but are acting toward a subverted goal; perhaps influenced to gain a lower privacy score as directed by the system. However, when looking at the "Yes" answers the red category is indeed the least disclosed with yellow being in the middle as one would expect and this also the case after salient review also.

This group was the lowest of all the treatment groups in terms of disclosure for both yes and no responses. Indeed, in terms of all responses made, there is a statistically significant decrease from both the PA and SN groups when compared to PC2 ($p=.044$ and .033 respectively). This would suggest that this treatment held the greatest sway of participant behaviour. However, a similar comparison when considering only yes responses does not yield statistically significant results (.079 and .068 respectively).

Given that disclosure also decreased dramatically in the green category of data, it is unclear whether changes in behaviour are as a result of an increased control of sensitive data or for other reasons. The extent of the levels of non-disclosure would suggest that the goal of interaction has been subverted from one of profile creation to a pure privacy oriented one, as mentioned earlier. In the post-experiment interviews one participant stated: *I deleted*

information to get my P-Score down. This would suggest that the interactive P-Score that dynamically changed when information was deleted provided a clear impetus to delete information and disclose less. The literature review chapter identified the goal-driven nature of users when interacting with systems (Jacko and Sears 2003) where users will act according to a perceived ultimate goal when dealing with sub-goals. This, perhaps, led to the control group disclosing much to achieve the perceived goal of account creation and disclosing little in this group to achieve the perceived goal of protecting one's privacy. Further evidence for this can be found in the connection settings. This group was the only group to make changes to these settings across numerous participants and this was the only group where these settings where explicitly put in front of the user with a dynamic score that encouraged the highest level of protection. This group also had the highest level privacy settings applied (although, again without statistical significance) with a standard deviation of 24.3 showing a much more consistent behaviour compared to other groups and providing evidence for the assertion that participants were encouraged to be highly private.

However, whether participants are acting toward their own goals or not is unclear. The design of the treatment was to aid in identifying and deleting potentially sensitive information; however, the level of non-disclosure in the green category of information would suggest that they are being persuaded to be more private than they perhaps need to or, indeed, desire to be. The privacy score and rating level (low-high risk assessment) may be too strict and participants may need to delete too much information to achieve the low risk assessment. Hence, future iterations of the experiment should reduce this level so that it is clear that green data will give a low risk and other data a higher rating.

8.3.1 Summary of Points

The groups with treatments present exhibited less disclosure with statistical significance. This decrease was present across the sensitivity ratings for all the groups but was the highest in the red sensitivity when examining only the "Yes" responses to questions. If awareness has been increased then one might expect disclosure to not lessen very much in the green category of data. However, this group of data items also saw decreases of similar levels to the other categories; furthermore, the "Yellow" category was the most answered in all the treatment groups suggesting a bias in the questions. It is, therefore, unclear if users chose to not disclose information based on their privacy preferences from this experiment alone.

Of note were the increases in standard deviation within the treatment groups which would suggest much more variation in participant behaviour leading to less consistent group levels of disclosure. This could demonstrate the individuality of privacy where, when reminded, some participants choose to protect themselves much more through non-disclosure across a variety of sensitivity ratings. Incidentally, the decreases across even non-sensitive categories does not necessarily mean that participants are not thinking about their privacy as a "green" item may be private to them. However, it is unclear here if this is the case due to the potential limitations in the experiment design summarised in a following section.

8.4.1 Exit Survey Results and Discussion

The first portion of the survey dealt with the Westin privacy rating typically used in measuring privacy concern; the goal of which was to examine the spread of concern across the groups and table 20, below, details the spread of the Westin ratings for each of the groups in this experiment.

Table 20 - Summary of Westin Ratings

Group	Fundamentalist (number/%)	Pragmatist (number/%)	Unconcerned (number/%)
Control	3 (30%)	6 (60%)	1 (10%)
PA	1 (9.1%)	5 (45%)	5 (45%)
SN	3 (25%)	9 (75%)	0 (0%)
PC	3 (27%)	8 (73%)	0 (0%)
Research	25%	57%	18%

Table 20 can only serve as an indication of the potential for salient influence due to the small sample size. It is interesting that the group with the highest number of fundamentalists is the control which also disclosed the most information. This may suggest that participants concern for their privacy was raised due to their realisation that they disclosed a great deal of personal information. The number of unconcerned was much increased in the personal attitude group; perhaps as participants had managed their privacy and were now not worried about it. This would not explain the decrease in the PC and SN group for this particular rating where no participant was unconcerned. They may have been influence by the treatments in question and adopted a more pragmatic view toward privacy; recall the PC group seemed to have been persuaded by the goal of privacy so may not have been representative of actual desires and the SN group may have been less in control of their behaviour as a result of conflicting information. However, the survey chapter also demonstrated that participants do not act

according to how they usually behave so this is may be an inaccurate measure and a larger sample size is required to examine if the spread of ratings are truly different in comparison to wider research.

The second phase of the survey was designed based on the surveys using the Theory of Planned Behaviour as a design tool. This gives three separate scores for behavioural intention, attitude, subjective norms and perceived control, the aim of which is to examine if the specific measure in the related group would be different from the others; i.e. will being exposed to a treatment based on perceived control give participants a greater perception of their level of control in the experiment. It is important to stress that a lack of increase in the respective measures does not mean that participants are not better equipped to deal with their privacy as survey measures have been shown to be inaccurate in reporting and explaining behaviour. The exit survey here is aimed at exploring this further; all scores are out of 7 for the following measures.

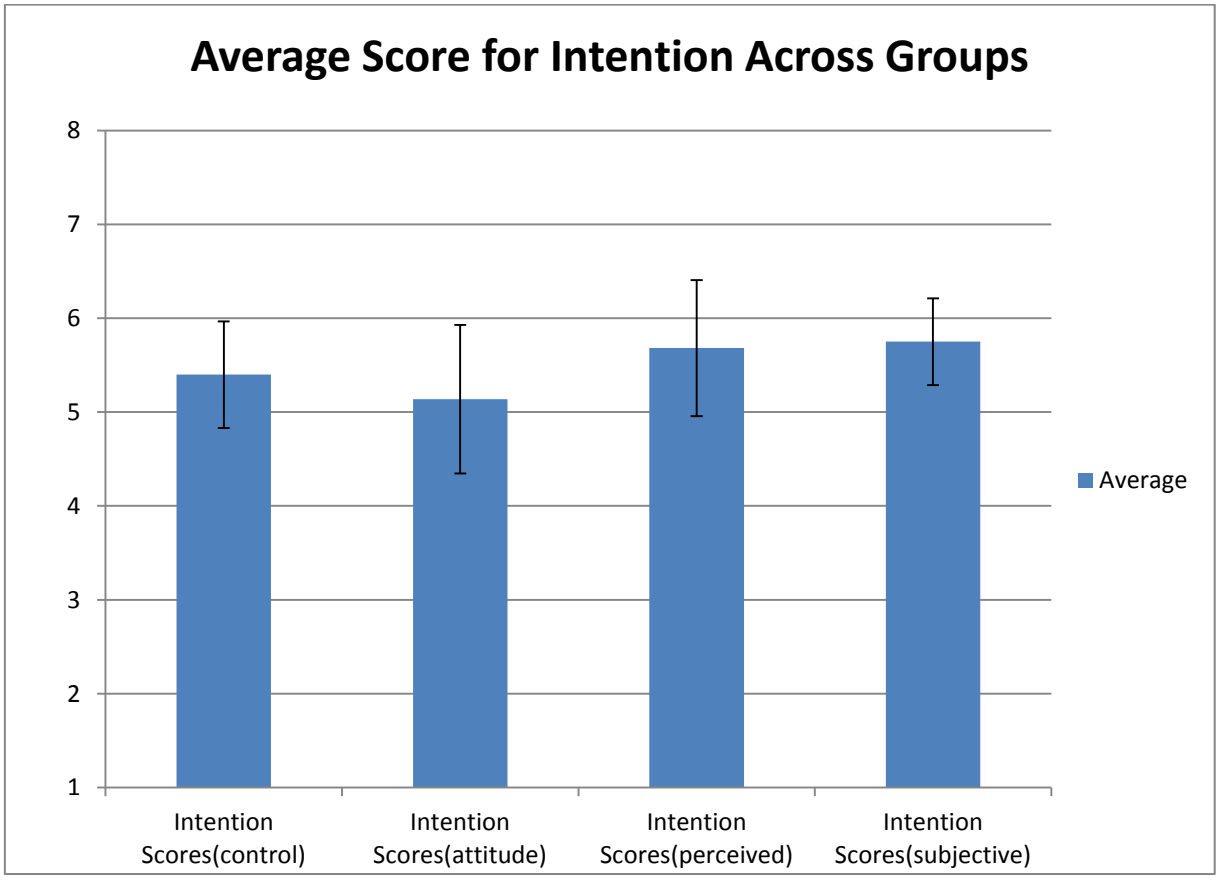


Figure 52 – Group Intention Scores

Figure 52 shows that the average score for intention was the lowest in the personal attitude group; a higher score would suggest a participant wishes to disclose as little as possible. This

is perhaps expected given that this group had the highest number of Westin unconcerned participants within it. The standard deviation was also the highest within this group demonstrating variation in the reported perception of participants within the group. The subjective norms group had the smallest standard deviation for the treatment groups and the highest average intention score despite having the highest amount of disclosure within the treatment groups.

Figure 53 details the spread of attitude scores across the groups with standard deviation bars added also.

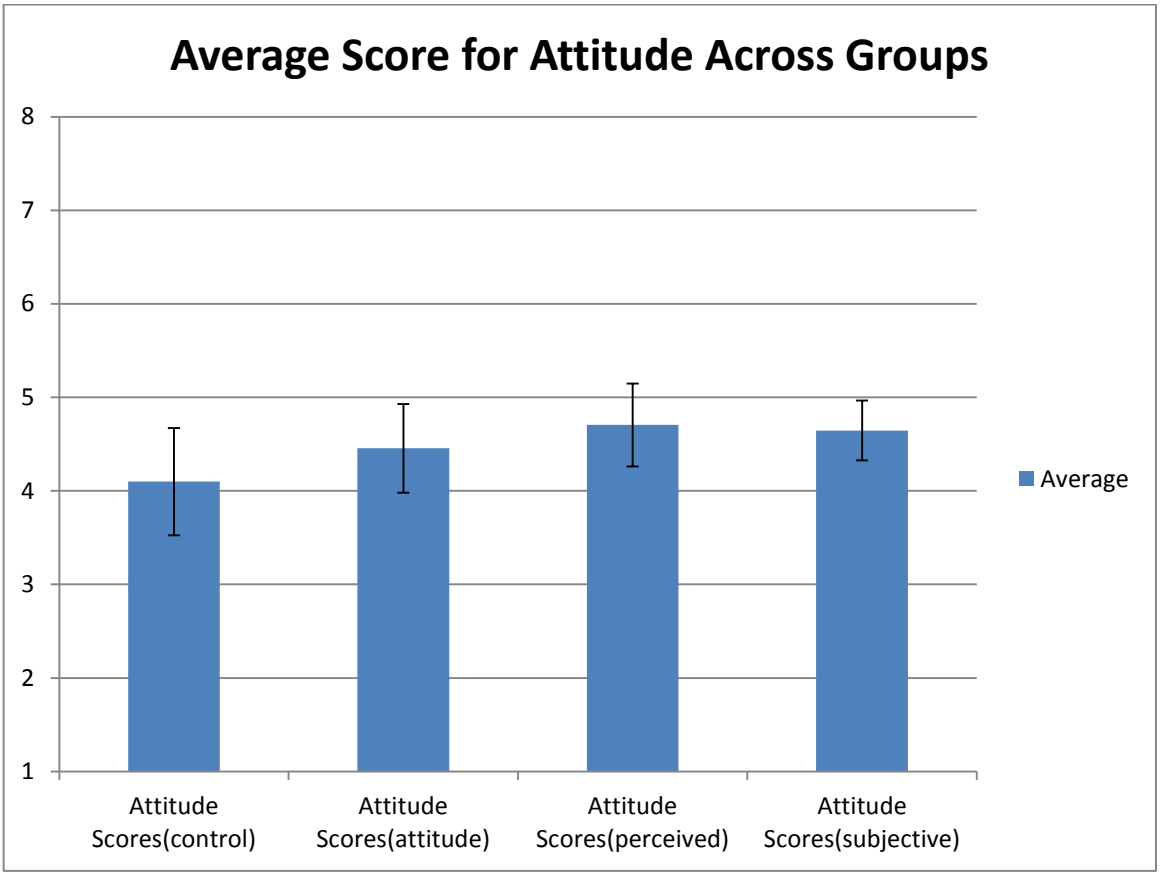


Figure 53 – Group Attitude Scores

A higher attitude score would suggests a positive attitude toward disclosure according to the TPB. The Control group participants reported the lowest attitude despite having the highest levels of disclosure. Again, this may be due to influence form the experiment making them more adverse to disclosure having realised they may have said too much. The spread across the treatment groups is broadly similar with the subjective norms group again having the smallest standard deviation demonstrating more consistent group responses. The attitude group did hold the smallest attitude score out of the treatment groups; however, this difference

is small. It is not therefore clear from this sample if participants have been influenced by the treatments and, indeed, the control group held the smallest attitude score. Again, surveys may be an inappropriate way of measuring perception and behaviour or the treatments could have the opposite effect of influencing the opposite direction; i.e. participant's perception of positive disclosure increased as they only disclose non-sensitive data. The average scores across the groups for reported perceived control scores are details in figure 54 below. A higher score here would suggest that participants believed that it was easy to identify and protect their potentially sensitive information.

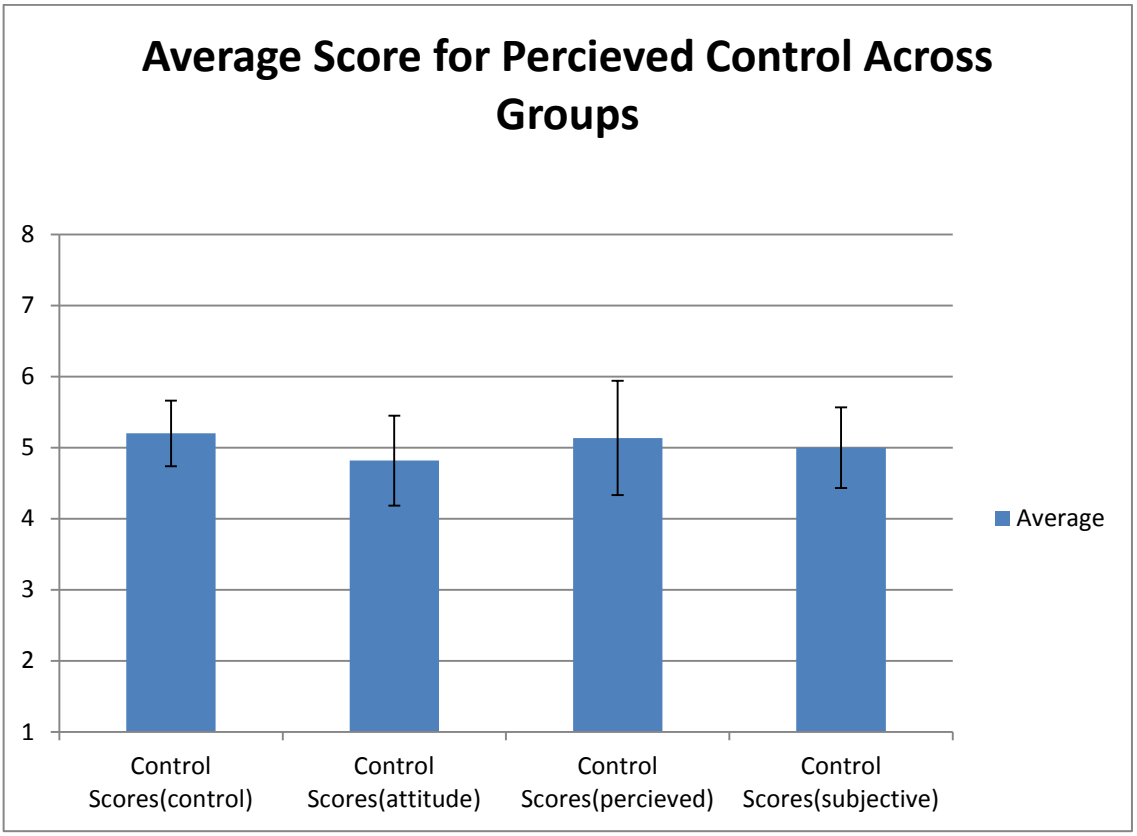


Figure 54 – Group PC Scores

A higher score would indicate that it was easy to identify and protect their sensitive information. The control held the (marginally) highest perceived control score for the groups, again, despite disclosing the most data. This group also held the most consistent responses as can be seen from the standard deviation bar. However, the scores are relatively even across the groups and differences could be down to variation rather than treatment influence. A larger sample may provide a clearer view of the relationship between perception and behaviour and the use of the TPB as a measure for each. However, from these results thus far it would appear that, much like in the survey previously, perception and behaviour cannot be

explored adequately through a general survey. This is a continuing theme in the measure for subjective norms (figure 55) where a higher score suggests that participants are more likely to perceive themselves to be susceptible to advice and social influence.

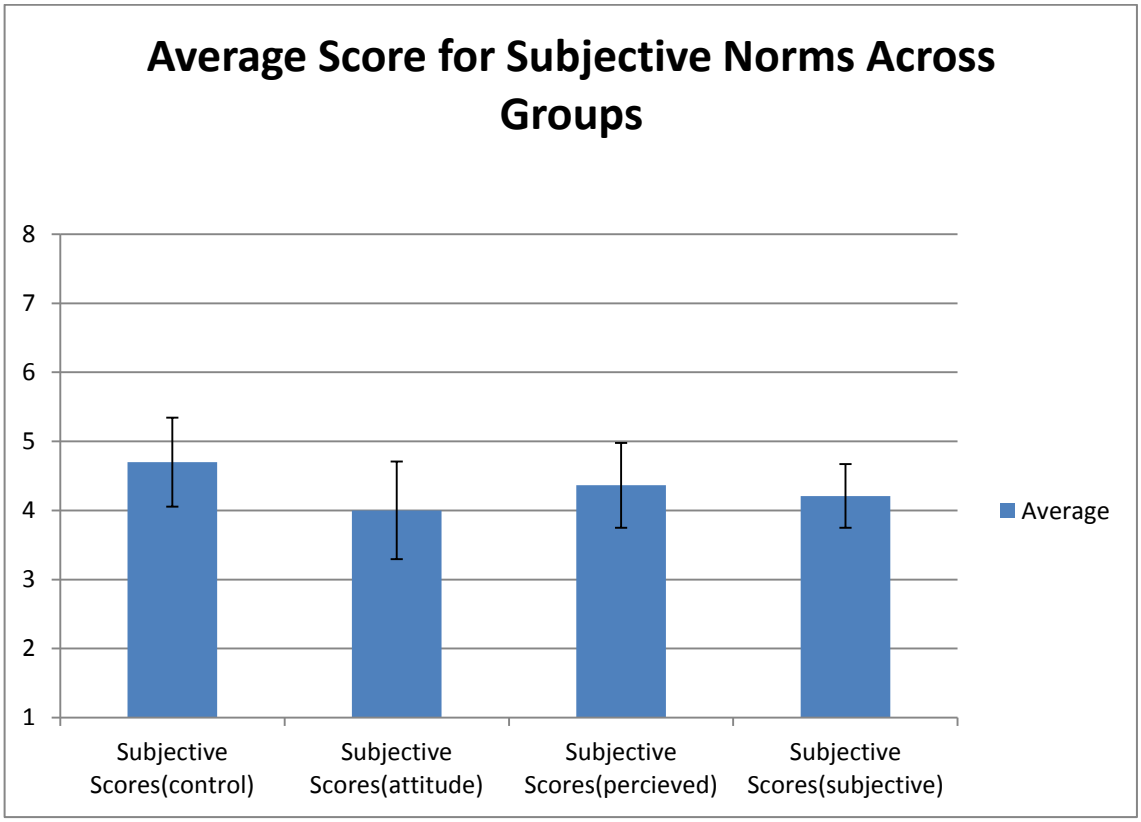


Figure 55 – Group SN Scores

A high score would indicate that they are likely to listen to others and again, the scores are relatively even across the groups with the control being the highest, as before. Despite their being very different observed behaviours between the groups the measure employed within this survey suggest minute differences in perception. The survey, as it is, does not seem capable of capturing the granularity and context of privacy such that it is relatable to actual behaviour. This may be one reason for the privacy paradox and this was touched on in the survey chapter also. That is, a general measure of perception cannot match the specifics of behaviour; for example, believing that disclosure can be beneficial may not indicate if a participant is unwilling to disclose their date of birth within specific context leading to a seemingly paradoxical observation. While, an automatic survey question could be generated to deal with what participants actually did in the experiment, the issues is one of information depth and a qualitative approach is required to explore these perception further. There is,

however, difficulty in implementing such an approach given that the data must be anonymised upon submission to protect the sensitive data being dealt with.

The surveys also offered no direct participant perceptions of the treatments themselves in terms of what they thought of them, whether they were a good thing, whether they believed themselves to be influenced by them or not etc.

8.5.1 Summary of Limitations and Recommendations for Change

As noted in the previous discussion there are limitations in the experiment design. First, it is unclear whether changes observed in treatment groups are a result of a rise in privacy awareness or due to a limitation in the design of the control group, namely a lack of a clear indication that disclosure is optional. The control group should be altered to make the optionality of disclosure clear so that subjects are actively choosing what to disclose based on their perceptions rather than any obligations they may feel.

Second, the questions asked may not be applicable to all participants leading to unanswered questions for reasons other than invasiveness; this may explain the decreases in the green category of information. Questions should be adjusted to provide an equal chance across all sensitivities for a response to be made. This is difficult as not all questions will apply to all participants (i.e. not all participant will read so cannot respond to favourite books). So, they must be general enough to elicit a response but sensitive enough for a tangible risk from that response.

Third, the treatment representing Subjective Norms seemed to confuse users by offering two types of conflicting advice. It is therefore unclear what the effect of the treatment itself is in terms of if participants are following one set of advice or the other or even if treatment at all is a persuasive UI element. Hence, the treatment should be simplified to provide only a single point of advice to ascertain if participants act upon it when interacting with the system.

Fourth, it is impossible to know precisely what the effect is of the treatments in this experiment as the relationships between the independent variables (treatments) are not explored. For example, the Perceived Control seemed to be driven by the goal of privacy protection (perhaps separate to their individually desired behaviours). Would this be the case if the Personal Attitude treatment was also present or would the rate of change after salient

review still be present? A factorial experiment design is required in order to explore this further and, ideally, should be implemented in a follow up experiment.

Fifth, the testing of H2, H4 and H6 is difficult given the lack of granularity available in exploring privacy settings. As participants seem to apply "Friends only" when they interact with these elements gaining statistically significant difference is difficult. Increasing the sample size or the variety of options available in this category may provide a solution.

Finally, the post-experiment interviews and exit-survey were insufficient in examining the behaviour from participants. The exit-surveys measurements of the TPB properties did not add to the explanation of the observations; this is perhaps due to the sample size or due to the general measures not dealing with the granularity that privacy requires as has been mentioned in the survey chapter. It is unclear whether the latter is the case, so the sample size should be increased and extra questions added to examine participant perceptions of the treatments and of their behaviour directly. Furthermore, focus groups should be added to explore the relationships that remain unclear; for example, whether participants prefer to lie or leave information out, whether one needs to apply protection where there is no disclosure etc.

8.5.2 Conclusions

This initial experiment has found a statistically significant difference between the treatment groups and the control group where the levels of disclosure are concerned. The settings scores, however, did not increase with statistical significance and were actually lower than the control in the Subjective Norms treatment group. The Perceived Control group held the lowest amount of disclosure compared to the other groups; perhaps due to the goal of interaction being very clearly oriented toward one of privacy. However, the decreases occurred within all sensitivity groupings and differences between them were negligible; although, the red category did have the highest percentage decrease from the green category within all groups, this was only when examining "Yes" answers. This could be due to participants disagreeing with the ratings of the questions asked of them and still implementing their own idea of privacy. However, it could also be due to limitations in the design of the experiment as outlined in the previous limitations section. In order to improve the value of the results, a further experiment was conducted taking into account the potential limitations and the following chapter details the specific changes, results and discussion for this second experiment.

Chapter 9 – Experiment Two

Chapter 9 – Experiment Two

9.1 Experiment Two Introduction

There is some evidence in experiment one for participants considering their privacy when salient information is present. However, due to limitations in the experimental design it is unclear if decreases in disclosure are due to other factors. As such, this chapter details a second experiment with modifications to the experiment design.

9.2.1 Design Changes

First, the control group potentially did not make the option of disclosure clear enough to participants meaning the decreases in disclosure in the treatment groups may not have been due to privacy related decision making. In order to provide this, two changes to the UI will be added; first, red asterisks (figure 56) will be added to the first page of the experiment process. These are to show which fields are required at the start and to demonstrate that the remaining fields are optional, these are added specifically because participants are used to seeing them in similar online forms as demonstrated by an interview response from experiment one.

Figure 56 – Making choice clear

On the profile builder page, a second notice is added to ensure it is clear on all pages (figure 57).

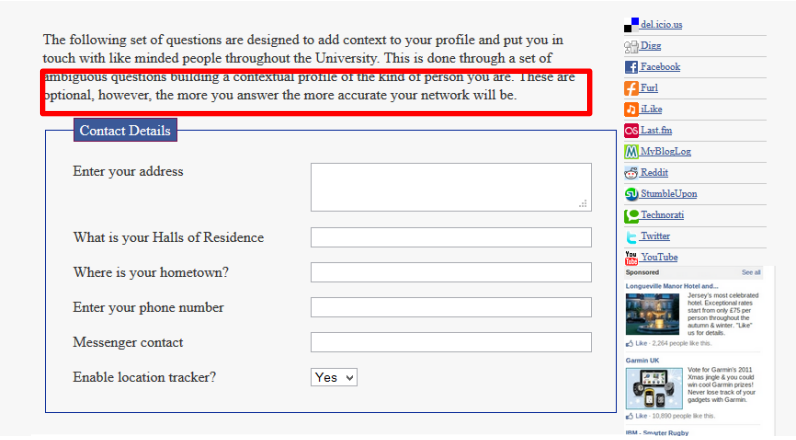


Figure 57 – Making choice clear

Note also, the Salford Network is changed to the NTU network as this experiment is conducted with students at Nottingham Trent University rather than the University of Salford. As such, convenience sampling is still to be used and the researcher has ready access to students at this University. This does itself raise some limitations which are to be discussed later in this chapter.

The second change deals with the question groupings. Responses in the green category may be due to participants simply not having an answer to the question leading to reduced disclosure in a low-sensitive category. Hence, some questions are altered in order to increase the chance they will be answered as they may be more applicable to more participants. Favourite quotes will therefore be changed to favourite TV Shows and Favourite Books will be changed to Favourite Music. Also, in order to compensate for the number of binary questions in the more sensitive question groupings, the green category will have two questions added asking users whether they are on the donor list and whether they donate to charity. The question dealing with downloading media explicitly says “illegally download…”. Two extra questions will be added to the yellow grouping to bring up the number also; this will ask the participant for their sexuality and personal email address. These are added as the yellow category in some groups in the previous experiment was answered highly, in some cases more so than the green category. Indeed, according to table 24 (below) the green category is more answered in experiment two compared to the other categories of information. Furthermore, the descriptors for the categories have been broadened to low, medium and high impact risks instead of social, professional and legal risks (although a statement is included to mention that they include these as potential, specific risks).

Furthermore, the settings have been adjusted, with interest, city and education being brought out of the connection settings and into the privacy settings. This is because these settings relate to content potentially being submitted during the experiment and thus may require protection. Furthermore, the photos and videos settings are altered to protect the contextual questions submitted with photos and videos being covered by status updates option. These profile settings are used to test H2, H4 and H6 and the connection link is kept only to see if participants explore it in privacy salient groups as a result of heightened concern.

The third change deals with the Subjective Norms treatment. It was noted that the conflicting information may have caused confusion for participants leading to unclear disclosure trends (across the groupings) and a decrease in privacy settings when compared to the control group. Hence, the advice containing information about what other users of the network are doing is removed leaving only the expert advice in the popup box. This expert advice advises against disclosing information in the yellow and red categories of information.

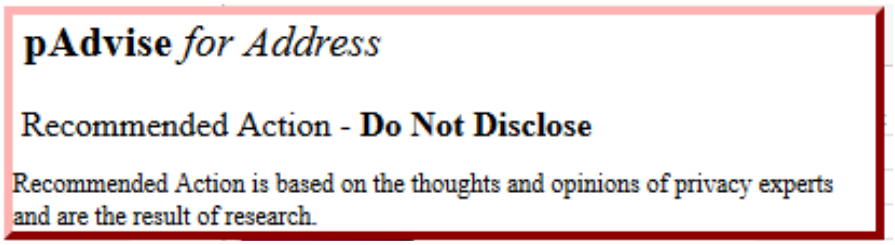


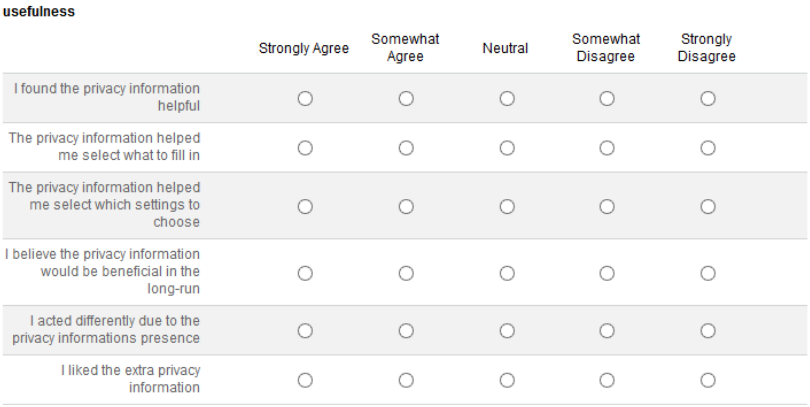
Figure 58 – SN single advice

Finally, three extra experiment groups will be added to change the experiment design to a factorial 2x2x2 design. This will involve a number of groups with combinations of the treatments presented to the user during the experiment. For example, one group's treatment will contain both the personal attitude and the subjective norms UI elements, one group with personal attitude and perceived control and so on. This gives the benefit of exploring the relationship between the treatments when implemented together. However, a full design will not be utilised (specifically there will not be a group containing all three treatments) as the UI in such a group may be too confusing and busy for the participants making it difficult to examine effect of the treatments; furthermore, there are limits in the number of participants available to include such a group.

9.2.2 Post-Experiment Changes

The exit–survey will have extra questions dealing directly with participant perceptions of the treatments they were exposed to. As such, all treatment groups will have an extra bank of

questions addressing how useful participants believed the UI elements to be and if they perceived themselves to be affected by them (figure 59).



usefulness

	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree
I found the privacy information helpful	○	○	○	○	○
The privacy information helped me select what to fill in	○	○	○	○	○
The privacy information helped me select which settings to choose	○	○	○	○	○
I believe the privacy information would be beneficial in the long-run	○	○	○	○	○
I acted differently due to the privacy informations presence	○	○	○	○	○
I liked the extra privacy information	○	○	○	○	○

Figure 59 – Added Exit-survey questions

Furthermore, the treatment groups based on factorial design will have added questions to ascertain which salient feature (if any) participants preferred as well as an extra two statement examining whether having two sources of information is confusing or beneficial. Each group has an added question asking participants whether they considered their privacy when answering questions during the experiment; this will be further explored in the qualitative interviews after survey participation.

The post-experiment interviews will take the form of focus groups in order to encourage discussion from participants. These will focus on the aspects of disclosure that cannot be satisfactorily explored through quantitative measures alone; namely, whether deception or withholding information is preferred from a user perspective and what the relationship is between disclosure and control (settings application). The focus groups will still be informal in nature and will address the following themes: Was your desire for privacy met? Did you think about your privacy when answering the questions? Did the salient features help you choose what to disclose and what was your perception of them? Is it better to leave out questions or answer "No"? Do you need to protect if you have not disclosed anything?

9.2.3 Procedure Changes

As mentioned previously, participants will be recruited from students of the School of Science & Technology at Nottingham Trent University, again being approached at the end of their scheduled lab sessions and asked if they would like to sign-up to a new social network just for NTU students (voluntary recruitment). Participants will be randomly allocated to a group and told not to discuss the process with others during it. The target sample size is

increased in order to provide greater power to the statistical tests ran; specifically, group sizes of around 20 are aimed for with the single treatment groups taking priority. This is to ensure sufficient numbers are obtained for these groups to allow for a thorough exploration of the singular treatments. The interviews will take place immediately after the experiment in order to maintain the context of the experiment and their perception of privacy within it as much as possible. These will be with around 2-3 participants from each of the groups involved.

The hypotheses being tested remain the same as found in chapter 7.

9.3.1 Initial Results – Single Factor Groups

In total 85 participants volunteered to take part in the single factor experiment and were randomly assigned to one of the seven groups; a breakdown of participant details can be found in appendix 10.

The total levels of disclosure are detailed in table 21 for all single factor groups in experiment two.

Table 21 - Disclosure summary for experiment 2

Group	Number of Participants	Average Total Amount of non-disclosure	Standard Deviation	Total % of questions answered	% of Yes answers when responded (total answered)
Control	20	5.8	7	82%	39% (201)
PA	23	11.3	4.6	66%	28.2% (174)
SN	21	14.2	6.8	57%	55.2% (125)
PC1	21	8.3	8.5	74%	62% (156)
PC2	21	16.7	9.6	48%	51% (95)

Of note here are the ratio of yes responses in the treatment groups SN and PC (pre-review) when compared to the control; although, the total amount of questions answered has decreased. It could be that participants are more likely to disclose if they have something to say and will leave it blank if not (given the higher ratio of positive responses).

Table 22 provides the same analysis with "No" answers counted as non-disclosure for the contextual binary questions.

Table 22 - Disclosure when only counting "Yes" responses

Group	Number of Participants	Average Total Amount of disclosure	Standard Deviation	Total % of questions answered
Control	20	11.9	5.9	63%
PA	23	16.9	4.2	50%
SN	21	16.4	5.6	49%
PC1	21	11.9	7.0	63%
PC2	21	18.9	7.9	41%

Table 23 details the results of statistical tests ran on the groups when comparing them the control. The non-parametric Mann Whitney U test is again utilised throughout the analysis of experiment two in this chapter as the data for some of the groups is not normally distributed for either levels of disclosure or settings scores.

Table 23 - Statistical Tests for levels of non-disclosure: treatments compared to control

Group	Test	P-Value (Total Questions answered)	P-Value (Discounting No Answers)
PA	Mann Whitney	<0.0001	=0.009
SN	Mann Whitney	<0.0001	=0.008
PC1	Mann Whitney	=0.272	=0.948
PC2	Mann Whitney	<0.0001	=0.003

An initial review would suggest that there is a statistically significant difference in behaviour between the groups, both with and without "yes" answers and that H1, H3 and H5's null hypotheses can be rejected. However, pre-review data within the perceived control group (PC1) did not decrease from the control group with statistical significance; this is unlike the initial experiment where both PC1 and PC2 held a statistically significant decrease when compared to the control. A difference may be expected given that participants in PC1 are exposed to the treatment before answering much of the questions asked of them. For example, a participant in PC1 will start at the welcome screen and answer the questions asked of them as in the control. Unlike the control, they will then review their question responses in the treatment screen before continuing with the experiment. Exposure to this review screen before visiting the proceeding question page (containing the majority of the questions) may influence their behaviour as they could be aware that they will again be asked to review their answers by the treatment.

Table 24 - Location of Disclosure across groupings (red highlights the least responded to category)

Group	% of "Green questions answered	% of "Yellow" questions answered	% of "Red" questions answered	% of "Yellow" questions, only yes	% of "Red" questions, only yes
Control	83%	82%	81%	60%	47%
PA	81%	66%	51%	48%	21%
SN	78%	50%	42%	40%	28%
PC1	77%	73%	73%	60%	52%
PC2	65%	41%	37%	32%	25%

Note there is no column in the table for "Only Yes" responses for the green category despite there now being yes/no questions in this category. This is as answering "No" to the green category yes/no questions is still considered an act of disclosure within this study and so counts as an answered question. For example, answering "No" to the green category question "Do you donate to charity?" reveals a potentially negative character trait about a participant, unlike answering "No" to the red category question "Have you ever downloaded illegal media?". Indeed, one participant stated that they felt it was worse to answer "No" to the green questions, so any response in the green category can be considered an act of potentially sensitive data disclosure.

Table 25 compared each of these groupings to the control to examine for statistically significant results.

Table 25 - Comparison of treatment categories to control categories

Group	Test	Green	Yellow	Red	Yellow, only yes	Red, only yes
PA	Mann Whitney	=.523	=.001	<.0001	=.005	<.0001
SN	Mann Whitney	=.311	<.0001	<.0001	=.004	=.004
PC1	Mann Whitney	=.242	=.192	=.175	=.937	=.317
PC2	Mann Whitney	=.027	<.0001	<.0001	=.001	=.008

Unlike experiment one, the decreases in the green category are not large and the major decreases did occur in the more sensitive yellow and red question groupings providing further evidence for H1, H3 and H5. This could be due to the changes in questions asked of participants that attempted to make them more applicable and more likely to be answered. The red category of questions hold the largest amount of percentage decreases from the control group (total amount of questions answered) which would initially suggest that participants are conscious of their privacy where the treatments are present; or at least persuaded by them to be more private. It is unclear whether they are implementing their privacy needs or those they perceive as required of them by the treatment (i.e. persuaded) and this will be explored further in the discussion.

The Perceived Control group, again, held the highest amount of non-disclosure when compared to the control; however, this is only after the review of data (PC2). Interestingly, the pre-salient review data is broadly similar to the control group. This may be expected given that the SNS aspects of the groups UI is not altered in any way from the control but does differ from behaviour observed in experiment one where the pre-review data was also significantly different from the control.

The average settings scores are detailed in table 26. These are now out of a total of 260 given the inclusion of the extra settings.

Table 26 - Average Settings per Participant

Group	Privacy Settings (St. Dev)	Connection Settings
Control	139 (116.5)	6
PA	147 (104.7)	2.6
SN	149 (113.2)	4
PC1	160 (95.3)	0
PC2	181 (92.7)	45.7 (53.4)

All groups held an on average increase from the control. Table 27 details the results of statistical tests ran on the above settings scores to examine if there is significant difference from the control group. Again, the Mann Whitney U test is used to assess significance.

Table 27 - Settings Statistical Tests

Group	P-Value
PA	P = 0.860
SN	P = 0.718
PC1	P = 0.543
PC2	P = 0.250

Much as in the first experiment, the changes in settings scores is not increased from the control with statistical significance. Participant perception of any potential relationship between these two variables is further explored through the post-experiment focus groups.

So, H2, H4 and H6's null hypotheses cannot be rejected within this experiment as there is no statistically significant difference when compared to the control much as in experiment one.

9.3.2 Initial Results – Factorial Groups

In total 42 participants were obtained for the factorial groups and split randomly between them. A summary of this group's behaviour can be found in table 28.

Table 28 - Factorial Groups Disclosure Overview

Group	Number of Participants	Average Total Amount of disclosure	Standard Deviation	Total % of questions answered	% of Yes answers when responded (total answered)
PA + SN	14	14	8.8	56%	31% (93)
SN + PC1	14	12.3	7.6	62%	39%(110)
SN + PC2	14	17.4	7	46%	34%(73)
PA + PC1	14	12.7	7.7	61%	49%(105)
PA + PC2	14	19.8	4.5	38%	57%(38)

Initially, it would appear that the results obtained are broadly similar to the single treatment groups that make up the relevant two factor treatments here. For example, the single factor SN group answered 57% of questions in total compared to SNPC1's 63%. Both the SNPC2 and PAPC2 scores (47% and 39%) are reduced compared to single factor PC2 (48%); albeit, only just in the case of SNPC2.

The levels of disclosure when only considering "Yes" responses are detailed in table 29.

Table 29 - Disclosure overview with only "Yes" responses

Group	Number of Participants	Average Total Amount of disclosure	Standard Deviation	Total % of questions answered
PA + SN	14	17.1	6.5	45%
SN + PC1	14	15.4	5.3	51%
SN + PC2	14	19.6	4.7	38%
PA + PC1	14	14.4	5.7	53%
PA + PC2	14	19.5	4.6	35%

Notice the difference between the two PC based groups is not as pronounced as the single factor treatments; again, as participants may have already made privacy choices after being exposed to the SN and PA treatments prior to review and resubmission.

Table 30 details the results of the test for statistical significance ran on the factorial groups in comparison to the control.

Table 30 - Factorial Disclosure Stats

Group	Test	P-Value (Total Questions answered)	P-Value (Discounting No Answers)
PA + SN	Mann Whitney	=0.004	=0.030
SN + PC1	Mann Whitney	=0.010	=0.033
SN + PC2	Mann Whitney	<0.0001	<0.0001
PA + PC1	Mann Whitney	=0.008	=0.158
PA + PC2	Mann Whitney	<0.0001	<0.0001

H1, H3 and H5 test true after an initial examination of these results when considering the culmination of the experiment (i.e. PC2). Unlike experiment one, groups including PC1 held reduced disclosure with statistical significance when compared to the control, perhaps as they have been exposed to the PA and SN treatment prior to reviewing their data. Interestingly, PAPC1 did not hold statistically significant difference compared to the control when only considering "Yes" responses, perhaps due to group variation. Table 31 breaks down the disclosure into the sensitivity categories.

Table 31 - Factorial Disclosure across categories (red highlights the least answered category)

Group	% of "Green questions answered	% of "Yellow" questions answered	% of "Red" questions answered	% of "Yellow" questions, only yes	% of "Red" questions, only yes
Control	83%	82%	81%	60%	47%
PA + SN	66%	58%	43%	46%	23%
SN + PC1	79%	61%	46%	48%	25%
SN + PC2	68%	40%	29%	32%	14%
PA + PC1	80%	55%	47%	46%	33%
PA + PC2	72%	23%	18%	20%	12%

Table 32 - Comparisons of treatment categories to control categories

Group	Test	Green	Yellow	Red	Yellow, only yes	Red, only yes
PA+SN	Mann Whitney	=.197	=.012	=.002	=.064	=.001
SN+PC1	Mann Whitney	=.569	=.027	=.012	=.138	=.004
SN+PC2	Mann Whitney	=.192	=.001	<.0001	=.003	<.0001
PA+PC1	Mann Whitney	=.500	=.007	=.009	=.071	=.061
PA+PC2	Mann Whitney	=.231	<.0001	<.0001	<.0001	<.0001

It would appear that the reduction in disclosure took place in the more sensitive question areas with the red category mainly having the least amount of questions answered. However, when considering only the "Yes" responses the reduction in the yellow category is not

significant when compared to the control (for initial submission in groups including the PC treatment). This is unlike the single factor treatments; it is unclear if this is due to group variation. The settings for the factorial groups are detailed in table 33.

Table 33 - Factorial Settings Summary

Group	Privacy Settings (St. Dev)	Connection Settings
PA + SN	156 (83)	10
SN + PC1	188.6 (107.2)	0
SN + PC2	240 (41.7)	57.9 (52.1)
PA + PC1	204.3 (98.2)	25
PA+PC2	228.6 (63.5)	76.4 (59.2)

The after review settings held higher on average increases than the single factor group of perceived control and the before review scores are also higher than the comparable single factor groups. This could be as their awareness of privacy was increased due to the combination of both the salient features present leading to a higher initial score that was then modified higher again.

Table 34 - Factorial Settings Stats

Group	P-Value
PA + SN	=0.931
SN + PC1	=0.306
SN + PC2	=0.017
PA + PC1	=0.110
PA+PC2	=0.025

Behaviour within the factorial groups containing perceived control did seem to hold statistical significance with P values < 0.05. There is, therefore, some evidence here for H6 given that upon salient review (PC2) there was statistically significant results but only when combined with PA and SN. This will be further explored within the discussion section

9.4.1 Discussion

The following discussion examines each group in turn to further explore the hypotheses presented in chapter 6 and will use the extra exit-survey data and participants responses in interviews to aid discussion.

9.4.2 Control Group

Table 24 would suggest that participants within the control did not consider the potential impact of disclosing sensitive information and therefore did not consider their privacy during the experiment. This is as disclosure was high across all sensitivity groupings; overall question responses were fairly consistent across the groups with 83%, 82% and 81% (60% and 47% for only "Yes" answers) response rates recorded.

When asked in the post-experiment review about whether they considered their privacy or not, participants responded: *I didn't think, having realised now, I think I would act differently*. This would appear to be a sound example of the privacy paradox at work and is in line with wider research where a similar response was recorded (Strater and Lipford 2008). As in experiment one, it may be that privacy is indeed a secondary goal problem and as wider work suggests (Bonneau, Anderson et al. 2009) and once reminded of it they wished they had considered it.

Despite participants being aware that disclosure was optional (indeed, all but three participants left something out) it was high in the group. In further exploration participants clarified: *It felt good to be able to fill in all the fields and complete my profile* and *I don't know why I answered the questions*. This would suggest that there is a sense of reward to being able to answer the questions asked of them; indeed, a study exploring disclosure and reward (Tamir and Mitchell 2012) found that the tendency to disclose information about the self is linked to the intrinsic value placed on it and that doing so releases dopamine making it a potentially addictive process. Hence, participants may wish to answer as much as possible due an ingrained desire driven by a subconscious process and the thought of privacy does not enter to the decision process unless reminded (as in the treatment groups).

For the settings, 57% of those participants who applied settings adopted an all-or-nothing approach and applied "Friends only" to all options and the connection settings were only set by one participant. This participant also held the highest amount of non-disclosure in the control group with 30 questions unanswered; suggesting that perhaps if there is a pre-defined notion of privacy concern, users will seek out risk-coping mechanisms as suggested by wider work (Youn 2009). When asked in the focus group if participants noticed the extra link they responded: *I didn't see it, I just went straight to the questions* and *I remember a link but it didn't register that it might be important*. Again, the concept of goal driven behaviour may

play an important role in the interaction with the experiment interface. The link was not noticed or given little thought as it was not seen as necessary to the process being engaged with. Furthermore, it would appear that promotion of increased system engagement is also required rather than just promoting awareness of the general concept of privacy; for example, where are the control features that would enable better privacy protection? Users must be enabled to find them as well as be encouraged to use them and, indeed, the Perceived Control group seeks to explore this further.

9.4.3 Personal Attitude

The Personal Attitude (PA) treatment was not modified from the previous experiment. Disclosure was less than the control with statistical significance, both in terms of total amount of questions answered and when only considering the "Yes" responses to binary questions. It is reasonable to expect disclosure to be lessened in the higher sensitivity areas (highlighted by the treatment) compared to the control as participants consider their privacy. Indeed, looking at table 24 would suggest that this is the case with 81% (green), 66% (yellow) and 51% (red) total questions answered across the three question groupings; of note, is that the amount of questions answered in the green category is not reduced by much compared to the control. When considering only "Yes" responses there is a dramatic difference in the yellow and red categories with 48% and 21%. However, whether or not disclosure is still occurring when answering is "No" is debateable and explored in greater depth in the focus group discussion later in the chapter.

Further evidence for participants within this group behaving in accordance to privacy concerns is provided through the exit-survey where 75% of participants in the group agreed with the statement "I acted differently due to the privacy information's presence".

Table 35 - Exit-Survey Summary - PA

Statement	Agreed	Neutral	Disagreed
I found the privacy information helpful	75%	19%	6%
The privacy information helped me select what to fill in	63%	25%	12%
The privacy information helped me select which settings to choose	75%	25%	0%
I believe the privacy information would be beneficial in the long-run	81%	19%	0%
I acted differently due to the privacy information's presence	75%	25%	0%
I liked the extra privacy information	75%	25%	0%

Post-experiment, a participant stated: *they did highlight ones that could cause problems, like address*. The focus on disclosure of Address potentially highlights the view of traditional privacy items highlighted in wider research (Stutzman 2006); although, this could have been used by the participant as a quick example. When asked if they agreed with the placement of the traffic lights (i.e. that the red items were the most sensitive etc.) the response included: *I could see why but I used that to make my own mind up* and *I made my decisions based on my own common sense*. In terms of H1, this would suggest that the treatment reminded participants of their own privacy desires rather than informing them in this particular case.

Interestingly, from table 35, 81% of participants felt the treatment would be beneficial in the long run of system use. This would suggest that these participants feel that this treatment may influence their decisions within such systems should their presence be a persistent one. Indeed, the focus group clarified: *it made me think twice about some of the information I put on Facebook*, this perhaps will give participants a privacy nudge as suggested by wider research (Wang, Leon et al. 2013).

H1 would appear to be well-founded based on results here. Participants disclosed significantly less than the control group and this disclosure was the least in the higher sensitivity ratings. Participant responses in the exit survey also show that a number of them perceived themselves to be affected by the treatment.

H2's null hypothesis cannot be rejected based on results here. Of those participants who filled in their settings, 44% selected "Friends Only" for all settings. One participant did select all settings that had an amber rating or higher but here is little evidence that this was the case throughout the group. However, 63% of participants did state that the treatment helped them fill in the settings so, while there is not statistically significant evidence that there was an effect, there is a large portion of users who perceived it to be useful based on the exit-survey results.

Furthermore, only one participant in the group chose to change any of the connection settings in the separate link. It would appear, therefore, that the UI treatment held some effect on participants decision of what to disclose but did not encourage greater engagement with the system itself. The treatments would seem to offer influence over clear interactive elements of the UI but not the more ambiguous aspects that would require a greater degree of desire to explore and understand; a driving reason to explore is needed from the user.

9.4.4 Subjective Norms

The Subjective Norms (SN) group held an advice pop-up with a single piece of expert direction telling the participants what to do for each question. This differed from the original experiment were there were two, often conflicting, pieces of advice making it unclear as to which the participant might be following (if any). For both the "Yellow" and "Red" categories of questions, the elements advice was to "not disclose" as the feature represents the advice of a privacy conscious "expert".

The group exhibited a statistically significant decrease compared to the control for both the total amount of questions answered and when only considering "Yes" responses to binary questions. The group also had lower disclosure than the personal attitude group; although, this is not a statistically significant decrease ($p=0.126$ for total questions answered and $p=0.887$ for only "Yes" responses). Much as in the previous groups the green category of questions were the most answered with a total of 78%, a small reduction of 5% from the control group for the same category (not statistically significant: see table 25). The more sensitive questions saw the largest reduction in question responses with the red category being the lowest. This would suggest that participants are considering their privacy during interaction with the system given the smaller reduction in the green category and the larger reduction in the other categories compared to the control.

The responses to the survey questions (table 36) are somewhat different to the PA group.

Table 36 - SN Exit Survey Responses

Statement	Agreed	Neutral	Disagreed
I found the privacy information helpful	72%	17%	11%
The privacy information helped me select what to fill in	56%	39%	5%
The privacy information helped me select which settings to choose	72%	22%	6%
I believe the privacy information would be beneficial in the long-run	83%	11%	6%
I acted differently due to the privacy information's presence	55%	11%	34%
I liked the extra privacy information	61%	39%	0%

Despite holding less disclosure than the control and PA group, fewer participants felt that the privacy information was useful or felt that they acted differently due to the treatments presence. In the focus group a participants stated: *I did think about my privacy because of it*

but I don't think I followed its advice. This may suggest that the advice offered by the treatment was too strict in its privacy assertion yet still made participants consider their privacy. However, the strength of the effect may further suggest that the persuasiveness of the property influenced participants to follow its advice despite their personal interests. Perhaps, the more the natural inclination to disclose is subverted the more participants resent being influenced.

Unlike the PA treatment, this UI element held a specific direction (rather than a general classification) that some participants may have found abrasive: *I thought it was too strict in general*. It is important to note that the focus group participants may not be representative of the whole sample. Furthermore, the majority of participants, according to the exit-survey, did feel that the UI feature was helpful (72%) and 83% felt that it would be beneficial in the long-run. Also, privacy is individually oriented, what may be private to one may not be to another so some of the yellow category questions may not have been sensitive for some individuals to disclose. However, the amount of disclosure within the yellow category was less than the PA group, perhaps as the treatment explicitly advises against disclosure as oppose to the moderate warning of a yellow light.

Much as in the previous group the settings scores were increased from the control on average but without statistical significance and, again, only one participant elected to change their privacy settings. It is worth noting, however, that 72% of participants felt that the treatment did help them select what settings to apply. Whether this provided justification for a choice they would make regardless (offering them with validation) or persuaded them when they were unclear is not apparent.

In summary, the group saw significant decrease in the levels of disclosure observed, particularly in the more sensitive categories of questions. However, participants within this group felt less effected by the treatment than in the PA group despite it actually having a stronger impact on their behaviour overall. Again, the settings were increased without statistical significance but participants did report finding the treatment a useful guide in choosing their privacy settings.

9.4.5 Perceived Control

The Perceived Control (PC) group allowed participants to review their information after submitting their responses. This review was intended to highlight sensitive information and give participants the chance to edit. In the previous experiment both PC1 (before review) and

PC2 (after review) saw a decrease in disclosure that held statistical significance. Here, however, only PC2 saw a statistically significant decrease from the control. PC1's observations are broadly similar to the control with only a 8% decrease in total questions answered. This could be expected given that PC1 is essentially the control interface with review screens added; i.e. the actual social network sections are the same as the control. The reason for the difference from the original experiment is unclear; however, the time between the two (variation in context) and variation in the sample used in each experiment could account for it. Furthermore, it could be due to the change in questions from experiment one where they are now designed to more applicable and therefore more likely to be answered by participants.

After review, both the total amount of questions answered and those including only the "Yes" responses were reduced with statistical significance with the green category of questions being the most answered (65%) and both the yellow and red holding the highest reduction in questions answered (41% and 37% respectively). This provides further evidence for H5 as participants would seem to be more selective and private with what they ultimately disclosed. This decrease when reviewing data is statistically significant with a Wilcoxon Signed Ranks score of $p<0.0001$ for both total questions answered and for only "Yes" responses. This is lower than the PA group also with a statistically significant difference ($p=0.042$) and lower than the SN group but without statistical significance ($p=0.385$). However, when only considering "Yes" responses there is no statistically significant difference compared to the PA and SN groups ($p=0.572$ and $p=0.442$ respectively). An initial review of these statistics would suggest that these participants are acting in accordance with strong privacy concerns for their information. However, this is only when reviewing their information so they may have been reminded of privacy when their interaction was placed in a privacy focussed context.

Table 37 - PC Exit-Survey Results

Statement	Agreed	Neutral	Disagreed
I found the privacy information helpful	58%	32%	10%
The privacy information helped me select what to fill in	63%	32%	5%
The privacy information helped me select which settings to choose	58%	37%	5%
I believe the privacy information would be beneficial in the long-run	42%	47%	12%
I acted differently due to the privacy information's presence	42%	37%	21%
I liked the extra privacy information	58%	32%	10%

Despite this group seeming to have the greatest effect on participant behaviour (largest decrease in disclosure compared to the control) the perception of how useful the privacy information/mechanism was lower than the other treatment groups detailed thus far. There appears to be a dissonance between accepting that a participant has acted differently and admitting to accepting aid. For example, 63% of participants agreed that the treatment was helpful; however, only 42% agreed that they acted differently due to it. It is also worth noting that every participant within this group modified their original submissions in some way so were, indeed, affected by the treatment. This highlights the difficulty in using self-reported perception to explore privacy that was originally identified in the survey chapter.

Statements post-experiment also suggest that the changes in the group are as a result of being driven by the dynamic elements added to the UI, particularly the changing "P-Score": *it was like a game, I wanted to get the high score*. It would therefore appear that participants may have been persuaded by the treatment to be more private but not necessarily to enact *their* privacy preferences; this persuasion by the treatment may be one way of overcoming the secondary goal problem privacy has been proposed as having (Bonneau and Anderson et al. 2009). However, the green category of questions did not see the same level of reduction as the other categories so it would appear that participants are still being selective of precisely what they are editing. The change in levels for the "P-Score" may have played a role in this as experiment one required a higher level of non-disclosure to obtain the "low risk" privacy rating.

This is the only group, much as in the first experiment, to set the connection scores. However, this is only upon review of their information when it is presented to them during the flow of interaction; that is, no participants followed the extra link before review. This would

suggest that participants will interact with what is put in front of them, particularly if there is sufficient motivation to do so as provided by the dynamic "P-Score". The settings score are also increased from the control but not with statistical significance.

In summary, this group held the largest decrease in disclosure when compared to the control and this is mainly in the more sensitive categories of information; although, the green also saw a reduction post-review. However, participants appeared to have been motivated by achieving a low "P-Score" rather than protecting their privacy. This is demonstrative of the power of persuasion using this type of UI element and there is evidence here that it can be used to force participants to act more privately. Although, if the goal is to encourage participants to enact their own privacy needs it is perhaps too persuasive; however, it is debatable as to whether participants will actually lose out if they are too private.

9.5.1 Two-Factor Groups

In order to explore the effect of the treatments further, this section discusses the results from the two-factor groups using focus group and exit-survey responses.

9.5.2 Subjective Norms & Personal Attitude

This group combined the SN and PA treatments so participants had both the traffic lights and an explicit piece of advice for each interaction. The advice and category definitions from the single factor treatments remain the same within the two-factor treatment found here.

Table 30 shows that both the total number of questions answered and disclosure where only "Yes" answers are counted held statistically significant reductions when compared to the control group. The total amount of questions answered stood at 56%, lower than the PA group at 66% and around the same level as the SN group which stood at 57%. Similarly, when "No" responses are discounted the groups level of disclosure stands at 48%. The location of this disclosure was least in the higher sensitivity ratings as in the single factor treatments from which this group was derived. The green category was also lower than the related single factor groups (66% compared to 83% and 78% respectively). There is no statistically significant difference between the 2-factor treatments and its respective single factor treatments ($p=.429$ for PA and $p=.829$ for SN in terms of total questions left unanswered)

It would appear that the two factor group held a similar affect over participants as the related single factor treatments and participants were making choices based on their privacy needs.

Table 38 - PASN Exit-survey results

Statement	Agreed	Neutral	Disagreed
I found the privacy information helpful	42%	42%	16%
The privacy information helped me select what to fill in	33%	58%	9%
The privacy information helped me select which settings to choose	33%	50%	17%
I believe the privacy information would be beneficial in the long-run	67%	25%	8%
I acted differently due to the privacy information's presence	42%	33%	25%
I liked the extra privacy information	67%	25%	8%
The presence of two sources of information was confusing	33%	58%	9%
The presence of two sources of information was better than one	36%	55%	9%

Participants here reported a more neutral view of the UI features than in the single factor groups despite being as affected by them, if not more so. In particular, participants reported that they did not feel affected by the treatments with only 33% agreeing with the statement regarding the helpfulness of it. However, 67% liked the privacy information in general and perhaps felt that it validated choices that they wanted to make.

Of the two salient UI changes 60% of participants preferred the Traffic Lights over the advice pop up, with the remaining participant stating neither had an impact. This is in line with the thoughts of the PA group who seemed to prefer their treatment over the SN group (when comparing the perceptions of the two). Indeed, the post experiment statements also highlighted this: *I preferred the lights, they were right next to each question so I was always aware of them even when not looking right at it* and when asked what they thought of the two at the same time: *I found the pop-up to be annoying making the site too busy*. There seems to be a general preference for the treatment introduced based on the Personal Attitude aspect of the Theory of Planned Behaviour; perhaps as it provided a suggestion rather than an order (Fogg 2003). The treatment may have allowed users to think about privacy without being too invasive in its intentions; i.e. not subverting their pre-defined goals as much allowing them to maintain a degree of autonomy over their own decisions.

The settings score was increased from the control but without statistical significance. It was, however, higher than the single factor groups barring PC2 which may suggest that participants held a heightened sense of privacy and made specific selections as a result. However, table 38 shows that only 33% of participants felt that the treatment helped them choose particular settings. Participants may have been subconsciously persuaded to act with a higher regard to their privacy (throughout the experiment) yet be unwilling to admit, or

unaware, that they were influenced by it. This may be similar to the third person effect were people tend to believe that mass communication plays a more significant influence over others than themselves and underplay the influence it actually has over the individual (Debatin, Lovejoy et al. 2009).

The connection settings were actually applied by three participants in this group. However, this could be due to variation within the sample rather than due to an increase in system engagement and participant self-efficacy given the lack of interaction in other groups.

9.5.3 Perceived Control and Personal Attitude

This treatment combined traffic lights of PA with the review pages of PC allowing participants to make initial selections of what to disclose based on the privacy salient information and then review and edit this information again. The levels of disclosure observed were least in the higher sensitivity ratings than in the green category suggesting participants were selective in terms of what they answered and either followed the advice of the treatment or regarded their own privacy during the experiment. The initial results in this group (before the PC treatment) are not statistically different from the single factor PA group suggesting similar behaviour ($p=.546$ and $p=.914$ for only "Yes" responses). It was also not statistically different from single factor PC1 ($p=.127$ and $p=.096$). Unlike single factor PC1 however, there was a statistically significant decrease when compared to the control (although, not when only considering "yes" responses). This would suggest that participants are making a selective decision over what to answer based on the PA treatment before submitting their data to the PC treatment. Having already made these choices would they then reconsider them based on the PC treatment? Disclosure was decreased post-review and the change in the group is significant for both yes and no responses (Wilcoxon $p=.028$ and $p=.001$) and PAPC2 was statistically reduced from the PA group (Mann Whitney $p<.0001$ for both counts of disclosure) suggesting that participants made their choices and were then persuaded to protect further by the control mechanism. However, despite there being a decrease compared to PC2 with a seemingly large reduction in the yellow and red categories (23% and 18% respectively) this is not statistically significant ($p=.175$ and $p=.325$).

Table 39 details the exit survey results.

Table 39 - PAPC Exit-Survey Results

Statement	Agreed	Neutral	Disagreed
I found the privacy information helpful	46%	38%	15%
The privacy information helped me select what to fill in	54%	23%	23%
The privacy information helped me select which settings to choose	38%	31%	31%
I believe the privacy information would be beneficial in the long-run	54%	31%	15%
I acted differently due to the privacy information's presence	38%	38%	23%
I liked the extra privacy information	46%	38%	15%
The presence of two sources of information was confusing	31%	46%	23%
The presence of two sources of information was better than one	23%	62%	15%

Much as in the PA and SN treatment group, participant responses are reduced in the agreed column from the single factor groups; despite the apparent effect of the treatment being greater than those groups. Participant's appreciation for the treatment seems to diminish the higher the potential effect of the treatment is. Again, perhaps being unwilling to admit the extent of its influence or finding the combination of the two an annoyance. Indeed, the focus group highlighted: *I'd already made my decision about what to disclose before reviewing, it was overkill*. In relation to this, 62% of participants in this group preferred the traffic lights over the review form (23% the review and 15% neither) suggesting that, for those participants, the added review may have provided an annoyance rather than an aid. So similar to the previous two factor group there appears to be a preference for the more suggestive traffic lights rather than the more explicit advice box. The reason for change despite already making privacy conscious decisions according to a participant echoes the response from the single PC group: *I wanted to get the score low.* Participants, despite potentially already making their privacy decisions, endeavoured to disclose less than perhaps they would have done otherwise and this seems to be due to the persuasiveness of the dynamic score giving them something tangible to aim for. Again, the influence of a sub-conscious goal may be playing a part in informing and influencing their behaviour as mentioned earlier; in this case the goal of getting a low score which may not be representative of actual participant desire. A Participant post-experiment felt that their choices were based on *their own common sense* rather than because of the treatment yet did not elaborate on whether they were reminded of their common sense due to the treatment of not. Indeed, they felt they may have been but could not say for certain.

Interestingly, the settings scores increased with statistical significance within this group when compared to the control; this and the other combined PC treatment were the only groups were this was the case. This statistical increase occurred upon the review of data after participant changed their originally submitted scores. Despite this, only 38% of participants agreed with the statement that the treatment helped them choose their settings and 31% outright disagreed with it. Furthermore, the group did not change the connection settings until reviewing their data when the settings are explicitly presented to the participant by the treatment; they may have felt that, when presented, they can choose the settings but that the treatment did not aid in that. Again, it maybe those participants are unaware of the effect that the treatment potentially held over them and felt that it did not play a role at all.

9.5.4 Perceived Control and Subjective Norms

This treatment provided participants with a pop-up advice box that they could follow while answering the questions and the PC review form to enable the modification and deletion of their responses. Disclosure in the group was statistically reduced when compared to the control with significant results. Compared to the single factor groups that make up this treatment the levels of disclosure is fairly similar with 62% prior to review (compared to 57% in the single factor SN) and 47% after review (compared to 48% in the single factor PC2). This similarity is also present when only considering "Yes" responses to questions. Indeed, there is no statistically significant difference to SN for either totals of disclosure (p=.516 and p=.960) nor is there compared to PC2 (p=.727 and p=.495) proving some confirmation of the results found in the single factor groups.

There is statistically significant difference post-review of submission for both yes and no counts of disclosure (Wilcoxon p=.002 and p=.002) and that final disclosure is not as pronounced as the previous PAPC group (however, no statistically significant difference: p=.376, p=.734).

This disclosure was more reduced in the yellow and red sensitivities of questions were participants were told not disclose information and the green seems to be fairly well answered at 79%. It would appear therefore, that participant did consider their privacy during the experiment and made their choices based on that consideration. The reduction in disclosure, upon participant review, is not as pronounced as the PA&PC group discussed previously perhaps as participants felt more confident in their responses due to the explicit order of the treatment SN. Indeed, participants did state post-experiment: *I got it right before review and*

did not really need it, but wanted to get the score down. This comment is very similar to a comment from the previous group showing some consistency of perception and the introduction of a clear privacy-related goal that the PC treatment provides.

The participant perceptions of the combined treatment can be found in table 40.

Table 40 - PCSN Exit Survey Results

Statement	Agreed	Neutral	Disagreed
I found the privacy information helpful	73%	23%	0%
The privacy information helped me select what to fill in	73%	23%	0%
The privacy information helped me select which settings to choose	64%	27%	9%
I believe the privacy information would be beneficial in the long-run	73%	18%	8%
I acted differently due to the privacy information's presence	45%	28%	27%
I liked the extra privacy information	55%	36%	9%
The presence of two sources of information was confusing	45%	55%	0%
The presence of two sources of information was better than one	45%	45%	10%

Unlike the previous combined treatment the participants in this group seem to have more favourable attitudes toward privacy information found within this experiment with less stated disagreement with the survey statements. However, it is unclear precisely which source of privacy information participants might be referring to in the combined treatments. However, 73% of participants felt that the treatments helped them choose what to fill in but only 45% felt that they acted differently due to its presence echoing findings in other groups. There is also no clear idea of which treatment participants preferred with 27% stating the review form and 27% the SN box (remained stating neither).

The settings scores in the group, after review, are statistically significant when compared to the control group (as in the other two factor group containing the PC treatment) and again the connection settings were applied upon presentation to the user on review. The standard deviation for the post-review settings stands at 42, the lowest of all the groups within the experiment suggesting fairly consistent behaviour across this treatment group and the majority of participant had settings at the highest possible score. This is an increased score with changes applied during the review of information. So, participants did make changes based on the salient information and mechanism provided by the treatment. Unlike the other

groups, 64% of participants agreed that the treatment did help to select their settings. This acceptance of the influence of the treatment is perhaps due to the treatment itself (direct order, clear behavioural reaction) or due to variation in the sample.

Again, the connection settings were only interacted with and altered upon review of data as seen in other treatments where PC mechanism is present. Hence, participants were not encouraged to explore the system to ensure their privacy needs were met and the treatments were review was not present did not aid in this. This could be due to their self-efficacy, so they were not confident in exploring the full features of the system or because exploring that link is simply not an obvious part of their goal and is therefore left alone and given low priority.

9.5.5 Focus Groups Discussion

The focus groups aimed to explore the more qualitative aspects of the experiments in greater detail. Specifically, examining participant perceptions on the link (if any) between disclosure and the application of privacy settings and whether there is a preference between lying or withholding information. Notes from these focus groups can be found in appendix 11.

First, it is unclear from these results what the relationship is between disclosure and privacy settings. Wider research suggests that there is no relationship (Christofides 2012) but an assumption could be made that if participants have disclosed very little they do not need to apply strong protection to their profiles. However, participants still applied high protection, even in the treatment groups that held less disclosure than the control. Indeed, one participant in the control focus group stated: *I chose out of habit, saw friends only and chose that*. This is an ongoing theme identified in the chapter 5 survey, where it seems that users of social networks, by default, want to apply the higher levels of settings protection and will do so when presented with them. The issue, perhaps, in social networks is that to change these settings requires effort on the part of the user that they are unwilling to give unless it is part of their regular interactions.

In response to the question "do you still need to protect even if you have not given any information", the response from all focus group participants was that it is necessary with on participant noting: *yes, things might change* and *you don't know what other people might say about you*. There is some evidence in these statements for participants thinking about the shifting context of privacy, acknowledging that the settings perhaps should be set to cover all eventualities that could occur in the future. It would appear therefore, that a relationship

between disclosure and settings should not be expected as they are each in pursuit of their own, different goals.

The second point the focus groups aimed to explore further dealt with whether participants prefer to withhold information or lie in response to questions asked of them in both the experiment and similar real world applications. The question was phrased thus: "Is it better to lie about your information or not answer the questions?" Particular focus was placed on the binary response questions. There was a general consensus within the focus groups that leaving questions unanswered is the better strategy when given the choice with one participant noting: *you are accountable for anything you put on there even if it is lie so just don't say anything*. Another participant stated: *if the data is being viewed then leave them blank as people will use that to judge you in some way*. This second response came from the PA group, the participant specifically stated that the lights helped show *what people may use to judge me on*. These views are in line with theories of social capital that suggest that any information about an individual can be used to form social ties and therefore, data should be withheld as much as is possible.

The data granule of address was chosen as a discussion point and when asked why participants did or did not disclose it a response was: *NTU does not need to know*. This was a common response to a variety of other data granules when asked. This would suggest that participants were making decisions based on what to disclose to the entity rather than their network. Participants also stated they held a similar approach to existing applications such as Facebook. However, it is important to note that this is only for information asked of them; i.e. the NTU network asked them these questions, not what they want to tell their friends using the system as a social mediator.

However, another participant stated: *It made me think carefully about the information I put on other social network sites such as Twitter and Facebook and how that information could be used by others.* This would suggest that the treatments did have the effect of informing participant awareness of privacy issues and them then using that awareness in their decision making process and hence may play a role in longer term use where disclosure is driven by user goal rather than the system goal.

9.6.1 Limitations

There is the potential that participants realised the purpose of the research and acted according to how they felt they were expected. Hence, participants chose not to disclose in the

higher sensitivities as they felt that was what was expected from their participation in an experiment rather than based on their own thoughts and perceptions on appropriate behaviour.

Also, the extent of the role of trust in the institution is unclear and could play a role in influencing the results of disclosure. That is, the amount of trust placed in NTU or Salford could mean participants disclose information that they would not normally do so to a similar service such as Facebook. However, a number of participants in the focus groups stated that they chose not to disclose certain pieces of information because *NTU does not need to know,* suggesting that participants were making a consideration based on the guardian of the data. However, the strength of the results may suggest that participants were enacting in accordance with perceived desirable behaviour.

The exit-survey highlighted that a number of participants felt that the privacy information provided by the treatments could be beneficial in the long-run. The experiment provided participants with a fixed context and a goal based on one that was given to them (to sign-up); hence, the experiment may not be representative of disclosure behaviour where the goal is derived by the user during long term system use. For example, a user may post on a timeline, not because Facebook asks them to, but because they wish to vent frustration or to inform their peer group of news. The exit-survey results suggest that participants may be reminded of their privacy if the information is present during longer term system use however.

As this is a post-test only control group design experiment, there is no gauge for pre-experiment behaviour from participants. Hence, any potential, pre-existing difference between the groups due to variation is unknown. A pre-test would be difficult to implement in these experiments as it may have the effect of priming participants to the idea of privacy and thus influence the results; again, Hawthorne effect. However, the size of effect on each group compared to the control (and the fact each group was different compared to the control) would suggest that any difference is caused by the treatment rather than variation.

Again, the exit-survey exploring differing perceptions of privacy concern and intention (based on Westin and the TPB) found no difference between the groups, much as in the first experiment; results can be found in appendix 10. This may be due to the survey not offering the level of granularity required to study the perception of discrete behaviour or due to the sample size not being sufficient in highlighting differences when using a survey as a data collection method.

9.8.1 Experiments Summary

Disclosure in the treatment groups was reduced when compared to the control sample with statistically significant results. Disclosure was the least in the more sensitive questions groupings with green seeing less decreases in disclosure compared to the first experiment. This could be due to the changes made to the questions were efforts were made to make the questions more applicable and therefore more likely to be answered by participants. However, the time between the two experiments could also be a factor and participants could be more engaged with the concept due to increased attention placed on privacy and the Internet (see PRISM (Greenwald and MacAskill 2013) etc.).

Regardless, as participants in this second experiment disclosed the least in the more sensitive questions groupings which would suggest that they were influenced by the treatments to consider their privacy. Hence, it would appear that H1, H3 and H5 tested true within this second experiment with participants reacting to the UI and disclosing less as a result. However, it is debatable as to the extent of influence the treatments held over participants in terms of meetings their own privacy needs. The Perceived Control groups, for example, seemed to be influenced by the dynamic mechanisms provided by the treatment and made their choices in pursuit of a goal of getting their P-Score as low as they could. Pursuit of this goal would provide the by-product of protecting a participant's privacy; however, participants cannot be said to be making choices based on their perceptions of privacy. Indeed, in the two factor groups, participants still made changes when reviewing their data despite already having made decisions based on the PA or SN treatment suggesting they were persuaded to be more private due to the PC treatment.

Participants seemed to favour the traffic lights when comparing perceptions across the groups (from the exit-surveys) and within the two-factor groups. This treatment is less tangible in it advice and is more of a general classification of question sensitivity: a suggestion. As such, it perhaps informs and reminds participants of their privacy desires while leaving them a degree of autonomy over their behaviour. Participants did follow the advice of the SN group but seemed less enamoured of the treatment compared to the PA group; perhaps not appreciating being told what to do.

The settings across the groups were not increased with statistical significance except for the in the two-factor PC groups were it was upon participant review. Hence, H2, H4 and H6 null hypotheses cannot be rejected for all groups. The settings are far less granular than the

disclosure levels and as such gaining statistically significant results is difficult based on the measures used.

Finally, the connection settings were only applied with consistency when presented to participants when reviewing their data and selections. This would suggest that engagement with the system was not heightened although engagement with the concept of privacy was increased. As in experiment one, the self-efficacy of participants when presented with a new system may be low and so they are unwilling to explore beyond the obvious bounds presented to them. Treatments applied here did not encourage this exploration and so participant privacy may have suffered as a result.

Chapter 10 – Conclusions

10.1 Conclusions

This chapter will review and answer the research questions posed in the introduction and literature review. It will review the approach taken within this study and propose recommendations for future iterations of such approaches within the research field. Finally, it will review the general contribution to the research field and propose further study possibilities based on the findings presented within this thesis.

10.2 Concluding the Method

This work has presented a method of exploring privacy behaviour and its relationship to the user interface using an experimental approach. This has proved a useful tool in examining whether there is a potential relationship between interface elements and what participants choose to answer. However, the settings application was less successful and perhaps is ill-suited to exploration using such an approach due to the lack of variation available in the measure. That is, participants tended to select an all-or-nothing approach to settings application and there are limited options available so measuring difference is difficult.

Such work requires a specific and difficult-to-design question bank to test levels of disclosure. These must be general enough to encourage answers from most participants to increase the likelihood that they have an answer to them (some participants may answer certain questions and other participants different questions). A consistent and fair question set is difficult to achieve within such a study. Despite this however, treatments groups can be compared to a control with the same question set to determine statistically significant differences.

A sub-aim of the experiment to assess self-efficacy through a separately kept connection settings was not successful and these were usually only set when directly presented to participants in the PC treatments. This came about as wider research suggests that participants are more likely to seek out risk-coping mechanisms when their concern is heightened (Youn 2009). However, this may be over a longer time period when the user is completely used to the system and aware of the wider risks of its use; indeed higher settings are associated with long term and active systems use (Lewis, Kaufman et al. 2008). Hence, the time-frame for exploring self-efficacy may be too short in this experiment or self-efficacy cannot be influenced in such a short time frame. It may be that prolonged exposure to the kinds of UI

elements described in this thesis may promote behaviour to seek out risk-coping mechanisms more.

The work deals with a controlled context of privacy and participants may act differently if they are outside of that context. For example, if users wish to post on their wall, they are directed under their own pre-defined goal instead of being asked an invasive question with a potential external goal. It is unclear whether such interface elements as those proposed within this thesis would have a similar, long term effect and be able to inform in such user-oriented situations. Certainly, some exit-survey responses suggest that there is some long term benefit to the privacy information added. As such, a longitudinal study would be appropriate where the same features are expanded into participants every day and less controlled routines. This perhaps could be achieved through Facebook's API or browser extensions. The features explored here could also form the basis for privacy recommender and personalisation systems.

The use of surveys based on general measures of intention and concern have been demonstrated to be ineffectual in exploring and predicting behaviour. The experiments included elements of surveys based on the TPB that were not granular or contextual enough to add to the dataset obtained. Perhaps a specific data item could be presented and participants asked for their perception of associated behaviour. However, the context would be lacking and difficult to recreate when testing if the perception was a predictor of behaviour. There is therefore a need for deeper qualitative enquiry to go along with the experiment. A full use of discourse and thematic analysis could reveal interesting participant perceptions about their behaviour within the experiment. For example, a phenomenographic analysis of participant responses post-interview could show that the groups with privacy salient information embedded hold deeper and more sophisticated qualitative responses to questions due to their exposure.

10.3 Impact and Contributions of the Work

The work has explored the concept of privacy behaviour from concern through to behaviour. The survey was designed to assess where the paradox occurs and to examine whether there is a relationship between concern, intention and behaviour finding that a disconnect occurs between users stated desires and their actual level of visibility in the network. This survey is also an exploration of the method in examining privacy and finds that it is perhaps not granular enough to examine the nuance of privacy. The Theory of Planned Behaviour is introduced and fitted to the causes of the poor privacy behaviour to demonstrate how they can

potentially be address by the UI. These were embedded into experiments aimed at exploring what happens when users interact with the system when there is a privacy salience in place, exploring varying types of reminder features. Findings indicate that where there are privacy focused mechanisms, participants will disclose less information about themselves when asked direct questions. In experiment two in particular, there appeared to be a focus on more sensitive questions where non-disclosure was concerned. This would suggest that users of social networks could be reminded or persuaded to be more private through simple UI features that introduce privacy salience. Ultimately, this work has made the following contributions:

1. A holistic exploration of the privacy paradox using a survey approach.
2. The Theory of Planned Behaviour is explored as a behavioural change mechanism for privacy behaviour.
3. Definitions of privacy salience are generated based on the TPB that address the causes of poor privacy behaviour as defined in literature.
4. A method is presented to explore the potential effect privacy salience might have on privacy behaviour.
5. This method is utilised and results presented of this effect within a controlled context.

This work has impact on HCI, user interface design and social psychology as well as making contributions toward privacy and technology work.

10.4 Directions for Further Research

The conclusions reached here raise some further questions for examining privacy behaviour in web services. First, do users of social networks have carefully crafted models of privacy to which they refer but which are subverted by the system they are interacting with, or are users simply victims of context and reacting to each individual situation they find themselves in? Secondly, is it simply the technology setting the context or are privacy preferences more stable in the offline world? Finally, would, therefore, the paradox be impossible to solve due to the instability of privacy preferences and the ease with which they are influenced by context?

The role of UI metaphors is also a recommended area for this work to move into and was touched upon briefly in the design of the salient features. Specifically, what are the potential metaphors for conceptualising privacy in such a way that system use can be quickly understood and privacy preferences easily implemented? Consider the real world, where privacy preferences are thought to be stable. The visualisation of privacy is obvious; closed doors and windows, thick walls, the ability to whisper etc. Such concepts are difficult to design through a system UI while maintaining the clarity of their intention. Technology which attempts to implement similar ideas such as encryption, tend to make the use of computer systems too complicated for most users to successfully utilise (Furnell, Jusoh et al. 2005). The goal then, is to examine ways in which these ideas can be modelled and implemented into a social network system and what effect would they have?

Building on this, how can salience evolve to become more effective? It is suggested that metaphors could also utilise tactile and audio based features making the user more productive (or in this case more private) (Marcus 1998). Hence, further questions could explore the effective or various interpretations of salience and examine there are varying degrees of influence depending on the sensory input.

A significant focus for this work was the actual interaction as it happens in real time. Observations during the course of the experiments were used to expand on the how users interact with the system but this is by no means all-encompassing or exhaustive. A similar study could be conducted using eye-tracking software to determine exactly what users are engaging with the most during interaction. Do they focus mostly on the actual form elements and ignore the surrounding UI elements? Do users in effect limit their contextual understanding of the system through this focus?

Finally, with increasing use of mobile devices the question is raised regarding how the recommendations put forward previously can be embedded within a more minimal UI of a mobile device. The previous point regarding a range of salient types could be incorporated with this point; sound in particular maybe worth exploring in conjunction with such devices. Furthermore, based on the discussions of context, are habits of disclosure different for mobile devices compared to their full system versions? This is an interesting thought, given that disclosure itself is easier when using a mobile system and that the context and setting within which the device is used could be far more changeable and also possibly play a role.

10.5 Concluding Remarks

This work has explored the potential effect of the User Interface on acts of behavioural disclosure by social networking users finding that the UI can play a role in producing more privately oriented behaviour through the use of cognitive models embedded within them. In a range of privacy contexts this could be an impossible task; however, the UI can persuade and subvert any other context as the final interaction point the user deals with. Whether or not the user's needs are being met in such circumstances is not answered and, indeed, it is debatable whether such a question can be answered given the complexity of privacy, the technology and users themselves. Hence, as designers of computer systems the fact that users misperceive system goals and their own goals must be taken into account and, ultimately one point must be kept in mind; that the only predictable aspect of end-users is that they are unpredictable. The Privacy Paradox phenomenon is the embodiment of such unpredictability; this thesis has proposed a way forward in designing UI's which allow for but limit the damage this unpredictable behaviour can potentially do.

References

(CommGAP), C. f. G. A. P. (2009). Theories of Behavioral Change. T. W. Bank.

Ackerman, M. S. and L. Cranor (1999). "Privacy Critics: UI Components to Safeguard Users' Privacy." Conference on Human Factors in Computing Systems: 258-259.

Ackerman, M. S., L. F. Cranor and J. Reagle (1999). "Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences." ACM Conference on Electronic Commerce: 1-8.

Ackerman, M. S. and S. D. Mainwaring (2005). Privacy Issues and Human-Computer Interaction. In L. Cranor & S. Garfinkel Security and Usability: Designing Secure Systems That People Can Use. L. Cranor and S. Garfinkel. Sebastopol, CA, O'Reilly**:** 381-400.

Acquisti, A. and R. Gross (2006). "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook." Proceedings of the 6th Workshop on Privacy Enhancing Technologies.

Acquisti, A. and J. Grossklags (2004). Privacy Attitudes and Privacy Behaviour: Losses, Gains and Hyperbolic Discounting. The Economics of Information Security. L. J. Camp and R. Lewis, Klewer.

Aimeur, E., S. Gambs and A. Ho (2009). "UPP: User Privacy Policy For Social Networking Sites." Fourth International Conference on Internet and Web Applications and Services: 267-272.

Ajzen, I. (1991). "The Theory of Planned Behaviour." Organizational Behaviour and Human Decision Processes **50**: 179-211.

Ajzen, I. and M. Fishbein (1980). Understanding Attitudes and Predicting Social Behaviour. Englewood Cliffs, NJ, Prentice Hall.

Bandimarte, M., A. Acquisti and G. Loewenstien (2012). Misplaced Confidences: Privacy and the Control Paradox. Workshop on the Economics of Information Security, Harvard.

Bandura, A. (1977). "Self-efficacy: Toward a Unifying Theory of Behavioural Change." Psychological Review **84**(2): 191-215.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ, US, Prentice Hall, Inc.

Barnes, S. B. (2006). "A Privacy Paradox: Social Networking in the United States." First Monday **11**(9).

Beck, L. and I. Ajzen (1991). "Predicting dishonest actions using the theory of planned behavior." Journal of Research in Personality **25**(3): 285-301.

Becker, J. and H. Chen (2009). "Measuring Privacy Risk on Online Social Networks." Web 2.0 Security and Privacy.

Binder, J., A. Howes and A. G. Sutcliffe (2009). "The Problem of conflicting social spheres: effects of network structure on experienced tension in social network sites." Proceedings of the 27th Internation Conference on Human Factors in Computing Systems. : 965-974.

Bishop, J. (2007). "Increasing participation in online communities: A framework for human–computer interaction." Computers in Human Behavior **23**: 1881-1893.

Bonneau, J., J. Anderson and L. Church (2009). "Privacy Suites: Shared Privacy for Social Networks." 5th Symposium on Usable Privacy and Security.

Boyd, D. M. and N. B. Ellison (2007). "Social Network Sites: Definition, History Scholarship." Journal of Computer Mediated Communication **13**(1): Article 11.

Brandimarte, M., A. Acquisti and G. Loewenstien (2012). Misplaced Confidences: Privacy and the Control Paradox. Workshop on the Economics of Information Security, Harvard.

Breakwell, G. M. (2006). Research Methods in Psychology. Oxford, Sage Publications Ltd.

Camp, L. J., C. McGrath and A. Genkina (2006). "Security and Morality: A Tale of User Deceit." In Models of Trust for the Web (MTW'06).

Card, S. K., T. P. Moran and A. Newell (1983). The Psychology of Human-Computer Interaction New Jersey, Lawrence Erlbaum Associates Inc.

Chen, S. and M.-A. Williams (2009). "Privacy in Social Networks: A Comparative Study." Pacific Asia Conference on Information Systems.

Chiasson, S., A. Forget, R. Biddle and P. C. van Oorschot (2008). Influencing Users Towards Better Passwords: Persuasive Cued Click-Points. Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction. Swinton. **1:** 121-130.

Christofides, E., A. Muise and S. Desmarais (2009). "Information Disclosure and Control on Facebook: Are They Two Sides of the Same Coin or Two Different Processes." CyberPsychology & Behaviour **12**(3): 341-345.

Church, L. (2008). "End User Security: The democratisation of security usability." 1st international workshop on Security and Human Behaviour.

Compeau, D. R. and C. A. Higgins (1995). "Application of Social Cognitive Theory to Training Computer Skills." Information Systems Research **6**(2): 118-143.

Consolvo, S., I. E. Smith, T. Matthews, A. LaMarca, J. Tabert and P. Powledge (2005). "Location Disclosure to Social Relations: Why, When & What People Want to Share." Conference on Human Factors in Computing Systems: 81-90.

Cozby, P. C. and S. C. Bates (2012). Methods in Behavioral Research. New York, McGraw Hill.

Cranor, L. F., P. Guduru and M. Arjula (2006). "User Interfaces for Privacy Agents." ACM transactions on Computer-Human Interaction **13**(2): 135-178.

Cranor, L. F., Reagle, J. and Ackerman, M. S. (2000). Beyond Concern: Understanding Net User's Attitudes About Online Privacy. Chapter in The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy. I. a. C. Vogelsang, B. M. United States of America, TPRC Inc.

Cranor, L. F. and J. R. Reidenberg (2002). "Can User Agents Accurately represent privacy notices." TPRC 2002.

Debatin, B., J. P. Lovejoy, A. K. Horn and B. N. Hughes (2009). "Facebook and online privacy: Attitudes, behaviors, and unintended consequences." Journal of Computer‐Mediated Communication **15**(1): 83-108.

Donath, J. and D. M. Boyd (2004). "Public Displays of Connection." BT Technology Journal **22**(4): 71-82.

Dourish, P. and K. Anderson (2006). "Collective Information Practice: Exploring Privacy and Security as a Social and Cultural Phenomena." Human-Computer Interaction **21**: 319-342.

Dwyer, C., S. R. Hiltz and K. Passerini (2007). "Trust and Privacy Concern within Social Networking Sites: A Comparison of Facebook and Myspace." Proceedings of the 13th Americas Conference on Information Systems

Fang, L. and K. LeFevre (2010). "Privacy Wizards for Social Networking Sites." World Wide Web Conference.

Fogel, J. and E. Nehmad (2008). "Internet Social Network Communities: Risk Taking, Trust and Privacy Concerns." Computers in Human Behavior **25**: 153-160.

Fogg, B. J. (1998). Persuasive Computers: Perspectives and Research Directions. Computer Human Interaction. Los Angeles**:** 225-232.

Fogg, B. J. (2003). Persuasive Technology: Using Computers to Change what We Think and Do. San Francisco, Morgan Kaufmann.

Fogg, B. J. (2009). The Behaviour Grid: 35 Ways Behaviour Can Change. PERSUASIVE. Clairemont, California.

Fogg, B. J. and D. Iizawa (2008). Online Persuasion in Facebook and Mixi: A Cross-Cultural Comparison. PERSUASIVE. O.-K. *e. al*. Berlin**:** 35-46.

Francis, J. J., M. P. Eccles, J. Marie, A. Walker, J. Grimshaw, R. Foy, E. F. S. Kaner, L. Smith and D. Bonetti (2004). Constructing Questionnaires Based on The Theory of Planned Behaviour: A Manual for Health Services Researchers. U. o. N. Center for Health Services Research.

Furnell, S. M., A. Jusoh and D. Katsabas (2005). "The Challenges of Understanding and using Security: A Survey of End-Users." Computers and Security **25**: 27-35.

Gable, G. (1994). "Integrating Case Study and Survey Research Methods: an Example from Information Systems." European Journal of Information Systems **3**(2): 112-126.

Galliers, R. J. (1991). Choosing Appropriate Information Systems Research Approaches: A Revised

Taxonomy. . Information Systems Research: Contemporary Approaches & Emergent Traditions. H.-E. K. Nissen, H. K. & Hirschheim, R. Amsterdam: North Holland.

Gandon, F. L. and N. M. Sadech (2004). "Semantic Web Technologies to Reconcile Privacy and Context Awareness." Web Semantics Journal **1**(3).

Goffman, E. (1959). The Presentation of Self in Everyday Life. New York, Doubleday.

Goh, K. N., Y. Y. Ghen, E. E. Mustapha, S. Sivapalan and S. Nordin (2009). Designing a Web Intervention to Change Youth Smoking Habits. HCI International J. A. Jacko. San Diego**:** 488-494.

Govani, T. and H. Pashley (2005). Student Awareness of the Privacy Implications When Using Facebook. Privacy Policy, Law and Technology Course, Carnegie Mellon University.

Grandison, T. and E. M. Maximilien (2008). "Towards Privacy Propagation in the Social Web." Web 2.0 Security and Privacy at the 2008 IEEE Symposium on Security and Privacy.

Greenwald, G. and E. MacAskill (2013). "NSA Prism program taps in to user data of Apple, Google and others." The Guardian **7**(6): 1-43.

Gross, R. and A. Acquisti (2005). "Information Revelation and Privacy in Online Social Networks." Workshop on Privacy in the Electronic Society.

Gross, R. and A. Acquisti (2005). Information Revelation and Privacy in Online Social Networks (The Facebook case). ACM Workshop on Privacy in the Electronic Society, Virginia.

Hampton, K. N., L. S. Goulet, L. Rainie and K. Purcell (2011). Social networking site and our lives. Pew Research Center's Internet & American Life Project. Washington, Pew Research Center.

Haslam, S. A. and C. McGarty (1998). Doing Psychology: An Introduction to Research Methods and Statistics. Wiltshire, Sage Publications Ltd.

Heeger, D. (1998). Signal Detection Theory. California.

Hochheiser, H. and J. Lazar (2007). "HCI and Societal Issues: a Framework for Engagement." International Journal of Human-Computer Studies **23**(3): 339-374.

Houghton, D. J. and A. Joinson (2010). "Privacy, Social Network Sites, and Social Relations." Journal of Technology in Human Services **28**(1-2): 74-94.

Hui, K.-L., H. H. Teo and S.-Y. T. Lee (2006). "The Value of Privacy Assurance: An Exploratory Field Experiment." MIS Quarterly **31**.

Iachello, G. and J. Hong (2007). "End-User Privacy in Human-Computer Interaction." Foundations and Trends in Human-Computer Interaction **1**(1): 1-137.

Jacko, J. A. and A. Sears (2003). The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications Mahwah, NJ, USA, Lawrence Erlbaum and Associates.

Jensen, C., C. Potts and C. Jensen (2005). "Privacy Practices of Internet Users: Self-Reports versus observed behaviour." International Journal Human-Computer Studies **63**(1-2): 203-227.

John, L. K., A. Acquisti and G. F. Leowenstein (2009). The Best of Strangers: Context Dependent Willingness to Divulge Personal Information. The Best of Strangers. Pittsburgh, Carnegie Mellon-University.

Johnson, J. P. (2010). "Targeted Advertising and Advertising Avoidance." Discussion Paper, Johnson Graduate School of Management, Cornell University.

Johnson, M., S. Egelman and S. M. Bellovin (2012). Facebook and privacy: it's complicated. Proceedings of the Eighth Symposium on Usable Privacy and Security. Washington, D.C., ACM**:** 1-15.

Johnstone, J., Eloff, J. H. P. and Labuschagne, L. (2003). "Security and Human Computer Interfaces." Computers and Security **22**(8): 675-684.

Joinson, A., C. Paine, T. Buchanan and U.-D. Reips (2007). "Measuring Self-Disclosure Online: blurring and non-repsonse to sensitive items in web-based surveys." Computers in Human Behavior.

Jones, H. and J. h. Soltren (2005). "Facebook: Threats to Privacy." from the Massachusetts Institute of Technology Web site: http://www.swiss.ai.mit.edu/6805/student.papers/fall05-papers/Facebook.pdf.

Kahneman, D. and A. Tversky (1979). "Prospect Theory: An analysis of Decision under Risk." Econometrica **47**(2): 263-291.

Knijnenburg, B. P., A. Kobsa and H. Jin (2013). "Dimensionality of information disclosure behavior." International Journal of Human-Computer Studies **71**(12): 1144-1162.

Kobsa, A. (2007). "Privacy-Enhanced Personalization." Communications of the ACM **50**(8): 24-33.

Kolter, J. and G. Pernul (2009). "Generating User-Understandable Privacy Preferences." International Conference on Availability, Reliability and Security: 299-306.

Krishnamurthy, B. and C. E. Wills (2008). "Characterizing Privacy in Online Social Networks." Workshop on Online Social Networks: 37-42.

Kumaraguru, P. and L. F. Cranor (2005). "Privacy Indexes: A Survey of Westin's Studies." Carnagie Mellon University, School of Computer Science, Pittsburgh: 22.

Lanheinrich, M. (2002). "Privacy Invasions in Ubiquitous Computing." workshop on socially-informed design of privacy enhancing solutions.

LaRose, R. and N. Rifon (2007). "Promoting i-Safety: effects of Privacy Warnings and Privacy Seals on Risk Assessment and Online Privacy Behaviour." The Journal of Consumer Affairs **41**(1): 127-149.

Lazar, J., J. Feng and H. Hochheiser (2010). Research Methods in Human Computer Interaction. Glasgow, John Wiley & Sons Ltd.

Lazar, J. and Y. Huang (2003). Designing Improved Error Messages for Web Browsers. Human Factors and Web Development. J. Ratner. USA, Laurance Erlbaum Associates.

Lederer, S., J. I. Hong, A. K. Dey and J. A. Landay (2004). "Personal Privacy Through Understanding and Action: Five Pitfalls for Designers." Personal and Ubiquitous Computing **8**(6): 440-141.

Lee, K. M., Y. Jung and C. Nass (2011). "Can User Choice Alter Experimental Findings in Human–Computer Interaction?: Similarity Attraction Versus Cognitive Dissonance in Social Responses to Synthetic Speech." Internation Journal in Human-Computer Interaction **27**(4): 307-322.

Legris, P., J. Ingham and P. Collerette (2003). "Why do people use information technology? A critical review of the technology acceptance mode." Information and Management **40**: 191-204.

Levi, M. and D. S. Wall (2004). "Technologies, Security and Privacy in the Post-9/11 European Information Society." Journal of Law and Society **31**(2): 194-220.

Lew, L., T. Nguyen, S. Messing and S. Westwood (2011). "Of Course I Wouldn't Do That in Real Life: Advancing the Arguments for Increasing Realism in HCI Experiments." Computer Human Interaction.

Lewis, K., J. Kaufman and N. Christakis (2008). "The taste for privacy: An analysis of college student privacy settings in an online social network." Journal of Computer - Mediated Communication **14**(1): 79-100.

Lipford, H. R., A. Besmer and J. Watson (2008). "Understanding Privacy Settings in Facebook with an Audience View." Proceedings of the 1st Conference on Usability, Psychology, and Security

Livingstone, S. (2008). "Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression." New Media and Society **10**(3): 393-411.

Lomax, R. G. (2007). "An introduction to statistical concepts."

Lyytined, K. (2010). "HCI Research: Future Directions that Matter " AIS Transactions on Human-Computer Interaction **2**(2): 22-25.

Macmillan, N. A. (2002). Signal Detection Theory. Stevens' Handbook of Experimental Psychology. J. Wixted. New York, John Wiley & Sons.

Marcus, A. (1998). "Metaphor design in User Interfaces." Journal of Computer Documentation **22**(2): 43-57.

Martiskainen, M. (2007). Affecting Consumer Behaviour on Energy Demand. SPRU - Science and Technology Policy Research. Sussex, University of Sussex.

Masiello, B. (2009). "Deconstructing the Privacy Experience." IEEE Security and Privacy **7**(4): 68-70.

Mathieson, K. (1991). "Prediciting User Intentions: Comparing the technology acceptance model with the theory of planned behaviour." Information Systems Research **2**(3): 173-191.

McConnaughy, E. A., J. O. Prochaska and W. F. Velicer (1983). "Stages of change in psychotherapy: Measurement and sample profiles." Psychotherapy: Theory, Research & Practice **20**(3): 368-375.

Milgram, S. and R. Fleissner (1974). Das Milgram-Experiment, Rowohlt.

Miller, R. E., M. Salmona and J. Melton (2011). Students and Social Networking Site: A Model of Inappropriate Posting. Proceedings of the Southern Association for Information Systems Conference, Atlanta.

Morgan, E. M., C. Snelson and P. R. Elison-Bowers (2010). "Image and Video Disclosure of Substance Use on Social Media Websites." Computers in Human Behavior **26**(6): 1405-1411.

Norberg, P. A., D. R. Horne and D. A. Horne (2007). "The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviours." The Journal of Consumer Affairs **41**(1): 100-126.

Oates, B. J. (2006). Researching Information Systems and Computing Oxford, Sage Publications Ltd.

Ott, L., M. Longnecker and R. L. Ott (2001). An introduction to statistical methods and data analysis, Duxbury Pacific Grove, CA.

Oulasvirta, A. (2008). Field Experiments in HCI: Promises and Challenges. Future Interactions Design: Part 2. P. Saariluoma, C. Roast and H. Punamaki, Springer.

Palen, L. and P. Dourish (2003). "Unpacking "Privacy" for a Networked World." Proceeding of the SIGCHI conference on Human Factors in Computer Systems.

Portes, A. (1998). "Social Capital: Its Origins and Applications in Modern Sociology." Annu. Rev. Sociol. **24**: 1-24.

Pötzsch, S. (2009). "Privacy Awareness: A Means to Solve the Privacy Paradox." The Future of Identity: 226-236.

Preibusch, S. (2010). "Experiments and formal methods for privacy research." Privacy and Usability Method pow-wow.

Quercia, D., R. Lambiotte, D. Stillwell, M. Kosinski and J. Crowcroft (2012). The Personality of Popular Facebook Users. ACM CSCW.

Rosenblum, D. (2007). "What Anyone Can Know: The Privacy Risks of Social Networking." IEEE Security and Privacy **5**(3): 40-49.

Russo, N. L. (2000). "Exploring the assumptions underlying information systems methodologies: Their impact on past, present and future ISM research." Information Technology and People **13**(4): 313-327.

Sasse, M. A., Ashenden, D., Lawrence, D., Coles-Kemp, L., Flechis, I. and Kearney, P. (2007). "Human Vulnerabilities in Security Systems: White Paper." Human Factors Working Group, Cyber Security KTN.

Sasse, M. A., S. Brostoff and D. Weirich (2001). "Transforming the 'Weakest Link' - a human/computer interaction approach to usable and effective security " BT Technology Journal **19**(3): 122-131.

Schneier, B. (2009). "Privacy Salience and Social Networking Sites." Schneier on Security - A blog covering security and security technology Retrieved July 16, 2009, from http://www.schneier.com/blog/archives/2009/07/privacy_salienc.html.

Sheppard, B. H., J. Hartwick and P. R. Warshaw (1988). "The Theory of Reasoned Action: A Meta-Analysis of Past Research with Recommendations for Modifications and Future Research." Journal of Consumer Research **15**: 325-343.

Smith, J. H., S. J. Milberg and S. J. Burke (1996). "Information Privacy: Measuring Individuals' Concerns about Organisational Practices." MIS Quarterly **20**(2): 167-196.

Smith, S. W. (2012). "Security and Cognitive Bias: Exploring the Role of the Mind." iEEE Security and Privacy **10**(5): 75-78.

Solove, D. J. (2006). "A Taxonomy of Privacy." University of Pennsylvania Law Review **154**(3): 477-557.

Solove, D. J. (2009). Understanding Privacy. Cambridge, Massachusetts, Harvard University Press.

Somekh, B. and C. Lewin (2009). Research Methods in the Social Sciences. Cornwall, Sage Publications Ltd.

Strater, K. and H. R. Lipford (2008). Strategies and Struggles with Privacy in an Online Social Networking Community. Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, British Computing Society.

Strater, K. and H. Richter (2007). "Examing Privacy and Disclosure in a Social Networking Community." Symposium on Usable Privacy and Security: 157-158.

Stutzman, F. (2006). "An Evaluation of Identity-Sharing Behavior in Social Network Communities." iDMAa Journal **3**(1).

Stutzman, F. and J. Kramer-Duffield (2010). "Friends Only: Examining a Privacy-Enhancing Behavior in Facebook." Computer Human Interaction.

Tamir, D. I. and J. P. Mitchell (2012). "Disclosing information about the self is intrinsically rewarding." Proceedings of the National Academy of Sciences **109**(21): 8038-8043.

Taylor, S. and P. A. Todd (1995). "Understanding Information Technology Usage: A Test of Competing Models." Information Systems Research **6**(2): 144-176.

Tonglet, M., P. S. Phillips and A. D. Read (2004). "Using the Theory of Planned Behaviour to Investigate the Determinants of Recycling Behaviour: a Casue Study from Brixworth, UK." Resources, Conservation and Recycling **41**: 191-214.

Tory, M. and T. Moller (2004). "Human Factors in Visualization Research." IEEE Transactions on Visualization and Computer Graphics **10**(1): 1-13.

Tuekci, Z. (2008). "Can You See Me Now? Audience and Disclosue Regulation in Online Social Network Sites." Bulletin of Science Technology Society **28**(20).

Turner, C. W., J. Nielsen and J. R. Lewis (2006). Determing Usability Test Sample Size. Internation Encyclopedia of Ergonomics and Human Factors. W. Karwowski. Boca Raton, CRC Press**:** 3084-3088.

Tuunainen, V. K., O. Pitkanen and M. Hovi (2009). "User Awareness of Privacy on Online Social Networking Sites - Case Facebook." 22nd Bled eConference.

Ugander, J., B. Karrer, L. Backstrom and C. Marlow (2011). "The Anatomy of the Facebook Social Graph." Arxiv preprint arXiv:1111.4503.

Valenzuela, S. (2009). "Is There Social Capital in a Social Network Site?: Facebook Use and College Students' Life Satisfaction, Trust, and Participation." Journal of Computer Mediated Communication **14**.

Wang, Y., P. G. Leon, K. Scott, X. Chen, A. Acquisti and L. F. Cranor (2013). Privacy nudges for social media: an exploratory Facebook study. Proceedings of the 22nd international conference on World Wide Web companion, International World Wide Web Conferences Steering Committee.

Webb, T. L. and P. Sheeran (2006). "Does Changing Behavioural Intentions Engender Behavior Change? A Meta-Analysis of the Experimental Evidence." Psychological Bulletin **132**(2): 249-268.

West, R. (2005). "Time for a change: putting the Transtheoretical (Stages of Change) Model to rest." Addiction **100**(8): 1036-1039.

Westin, A. and H. Interactive (1999). IBM-Harris Multi-National Consumer Privacy Survey, Tech. rep., 1999. Approximately 5,000 adults of the US, Britain and Germany.

Westin, A. F. (1991). "Harris-Equifax consumer privacy survey 1991." Atlanta, GA: Equifax Inc.

Westin, A. F. (2003). "Social and Political Dimensions of Privacy " Journal of Social Sciences **59**(2): 431-453.

Westin, A. F., D. Maurici, L. Price Waterhouse and L. Harris (1998). E-commerce & privacy: What net users want, Privacy & American Business Hackensack, NJ.

Whalen, T. and K. M. Inkpen (2005). Gathering evidence: use of visual security cues in web browsers. Graphics Interface, Victoria, British Columbia.

Woo, J. (2006). "The right not to be identified: privacy and anonymity in the interactive media environment." New Media and Society **8**(6): 949-967.

Wu, M., R. C. Miller and S. Garfinkel (2006). Do Security Toolbars Actually Prevent Phishing Attacks. CHI, Montreal, Quebec, ACM.

Youn, S. (2009). "Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents." Journal of Consumer Affairs **43**(3): 389-418.

Appendices

Contents

Appendix 1 – Privacy Perceptions Survey

Dear Student,

As part of my PhD on Data privacy in Web Based Services I am conducting research into user perceptions of privacy at a granular level and how their understanding of this affects their behaviour and therefore their own end data privacy. ***Granular Privacy means the individual fields and parameters (e.g. Date of Birth, address etc.) which relate to the users of a web service; in particular a Social Networking System.***

I am writing to ask for your consent to take part in this **questionnaire survey about your thoughts on the information held about you should you use a Social Networking System such as Facebook**. All information received and recorded through the use of this survey shall be used for no other purpose other than that related to the research and any related, possible publications. All data will be stored safely and confidentially during the study.

The research study is set to investigate user perceptions of privacy at the granular level and how private they believe the granular parameters to be and compare this to relevant action with regard to protection behaviour. Your contribution will be vital in examining the idea of privacy from a new perspective and enable myself and other researchers to develop privacy protection mechanisms which provide holistic and contextual aid to the user when deciding on appropriate privacy safeguards. It will demonstrate the complexity and individuality of privacy; also aiding in the understanding of what users actually think about ***their own*** privacy helping to provide the evidence for the need of individually tailored tools.

Some participants may be selected for follow up studies which include; an expert analysis of 10% of the participant's social network settings (this will be entirely anonymous and absolutely no personal data shall be recorded) and a follow-up focus group which explore further the themes of this survey. If you are happy to take possible take part in these further studies please make this clear on the next page.

At any stage of the research **you have the rights to withdraw your information.** Throughout the study **you also hold the complete rights to ask any further questions regarding the study which occur to you.**

When the study is concluded **you will be given access to a summary of the findings of the study and these shall be made available on request where publication has not been possible.**

If this is agreeable to you, please sign the attached research consent form.

I would like to take this opportunity to thank you for taking the time to help me with my research. If you have any queries, please feel free to contact me at the following email; T.Hughes-Roberts@edu.salford.ac.uk

Yours Sincerely,

Thomas Hughes-Roberts
PhD Student
The Computer Networking and Telecommunications Research Group
School of Computing, Science and Engineering
University of Salford

User Awareness/concern of Privacy Granules in Social Networks and Corresponding Behaviour for Protection

Principal Researcher: Thomas Hughes-Roberts

School of Computer Science and Engineering

1. I confirm that I have read and understand the information provided for the above study. I have had the opportunity to consider the information, ask questions and have these answered satisfactorily. ☐

2. I understand that my participation is entirely voluntary and that I am free to withdraw at any time, without giving reason and that this will not affect my legal rights. ☐
3. I understand that any personal information collected during the study shall be kept entirely confidential. ☐

4. I agree to take part in the study. ☐

5. I would be willing to take part in any relevant follow-up research regarding this project and am happy to be contacted regarding this. ☐

Researcher Name: Thomas Hughes-Roberts **Signature:**	**Date:**
Participant Print Name: **Email:** **Signature:**	**Date:**

Appendix 2 - Privacy Awareness Survey

Section 1 – Participant Information

Are you Male or Female?

A. ☐ Male
B. ☐ Female

Age:

A. ☐ 18 – 21
B. ☐ 22 – 25
C. ☐ 26 – 28
D. ☐ 29 – 32
E. ☐ 33+

Please specify your home country;

Do you actively use a Social Networking Site?

A. ☐ Yes
B. ☐ No

If no, please only complete sections 2 and 4.

Section 2 – Privacy Opinion and Concern Assessment

For the following statements please select the option which most applies to you by placing a tick in the appropriate box;

1. Users have lost all control over how their personal information is collected and used by social networking sites.

Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree

2. Social Networking sites handle personal information they collect in a proper and confidential way.

Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree

3. Existing laws and site policies/practices provide a reasonable level of protection for user privacy today.

Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree

Concern

1. How concerned are you about your personal privacy when using a social networking site?

Very Concerned	Somewhat Concerned	Not very concerned	No concern	Do not know	Refused

2. Have you ever been the victim of a perceived invasion of your personal privacy?

Yes, myself	No, but someone I know	Not at all	Do not know

Section 3 – Social Network Use

Please write a brief sentence on what you believe privacy means to you;

How regularly do you use a Social Networking Site?

E. ☐ Many times a day
F. ☐ Once a day
G. ☐ Many times a week
H. ☐ Less than once a week

Have you read the privacy policy related to the Social Network System?

C. ☐ Yes
D. ☐ No

Why do you use it (select as many as apply)?

J. ☐ Keep in touch with friends
K. ☐ Keep in touch with colleagues
L. ☐ Get to know new people
M. ☐ Easily obtain information regarding work/university
N. ☐ Show information about myself/advertise
O. ☐ Make it convenient for people to get in touch with me
P. ☐ Build relationships
Q. ☐ Find Jobs
R. ☐ Other, please specify;

How many friends do you have listed in the Social Network System?

F. ☐ 0-50
G. ☐ 50-100
H. ☐ 100-200
I. ☐ 200-400
J. ☐ 400+

What type of people do you add as a friend on Social Networks (select all that apply)?

I. ☐ Close friends
J. ☐ Family members
K. ☐ Friends you may not consider close
L. ☐ Colleagues you may not consider friends
M. ☐ People you know but do not consider friends
N. ☐ People you have met but once
O. ☐ People you have never met
P. ☐ Other, please specify;

Do you use the "custom" feature to group you friends list into types of people?

C. ☐ Yes
D. ☐ No

If no, why not (select all that apply)?

F. ☐ Unaware of ability to do so
G. ☐ Aware but do not know how
H. ☐ Do not want to utilize feature
I. ☐ Too time consuming to do so
J. ☐ Other, please specify;

What would a person not on your friends list be able to see do you believe (select all that apply)?

H. ☐ My Friends
I. ☐ My Groups/Networks
J. ☐ My Info
K. ☐ My Pages
L. ☐ My Photos
M. ☐ My Wall
N. ☐ Do not know

Section 4 – Granular Privacy Perception

Please rate the following pieces of information according to how likely it is that you would share that information with the following groups (1 being not likely at all and 5 being very likely).

Directly Personal Information

Full Name	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Date of Birth	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Address	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Email and Messenger	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Phone Number	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Home Town	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Current Location	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Family Members	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Relationship details	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Activities	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Favourites (films, books)	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Religious & political views	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Universities & Schools	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Work Place	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Sexuality	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Information posted by others

Photos by others	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Videos by others	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Posts & Comments by others	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Information posted by yourself

Wall Posts & Comments	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Status Updates	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Photo albums	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Videos	**Not Likely**				**Very Likely**
Strangers	☐1	☐2	☐3	☐4	☐5
Work/School Colleagues	☐1	☐2	☐3	☐4	☐5
Close Friends	☐1	☐2	☐3	☐4	☐5

Section 5 – Privacy Settings you have

The following are a list of the privacy controls available in Facebook. For each please select which you use to protect your information. If you are unsure at all please select "Not Sure"; do not guess which one you believe to be correct.

Please Note; FOAF means Friend of a Friend.

Factors	**How do you protect your information?**					
Primary Information	Everyone	Friends and Networks	FOAF	Only Friends	Custom	Not Sure
Full Name						
Date of Birth						
Address						
Email address						
Phone Numbers						
Home Town						
Instant Messenger						
Contextual Information	Everyone	Friends and Networks	FOAF	Only Friends	Custom	Not Sure
Personal Info (activities, etc.)						
Religious & Political Views						
Education and Work						
About Me						
Family & Relationships						
Unintentional Disclosure	Everyone	Friends and Networks	FOAF	Only Friends	Custom	Not Sure
Photos and Videos of me						
Who can comment on your profile?						
Posts by Others						
Intended Disclosure	Everyone	Friends and Networks	FOAF	Only Friends	Custom	Not Sure
Posts by me (links, updates etc.)						
Photo Albums						

Section 6 – Policy Awareness

Are you aware that Facebook owns any information uploaded into the site (i.e. are you aware that it owns your information)?

A. ☐ Yes
B. ☐ No

If yes, does this affect your behaviour on the site (i.e. are you less likely to put certain information on there)?

A. ☐ Yes
B. ☐ No
C. ☐ Do not know

If no, will you now modify your behaviour (i.e. are you less likely to put certain information on there)?

A. ☐ Yes
B. ☐ No
C. ☐ Do not know

Finally, select which option most applies to you; I am very concerned that Facebook owns the information I give it and can do with it what they want.

Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree

Appendix 3 – Survey Results and charts

Overview and Breakdown of Sample

Degree * Gender Crosstabulation

			Gender		
			Male	Female	Total
Degree	**Nursing**	Count	11	50	61
		% within Degree	18.0%	82.0%	100.0%
		% of Total	3.2%	14.7%	17.9%
	Spanish	Count	9	25	34
		% within Degree	26.5%	73.5%	100.0%
		% of Total	2.6%	7.4%	10.0%
	Law	Count	17	53	70
		% within Degree	24.3%	75.7%	100.0%
		% of Total	5.0%	15.6%	20.6%
	English	Count	13	43	56
		% within Degree	23.2%	76.8%	100.0%
		% of Total	3.8%	12.6%	16.5%
	DipN	Count	12	36	48
		% within Degree	25.0%	75.0%	100.0%
		% of Total	3.5%	10.6%	14.1%
	DipN2	Count	8	63	71
		% within Degree	11.3%	88.7%	100.0%
		% of Total	2.4%	18.5%	20.9%
Total		Count	70	270	340
		% within Degree	20.6%	79.4%	100.0%
		% of Total	20.6%	79.4%	100.0%

DipN and DipN2 include students from the same module studied; however, data collection took place over two days with DipN on the first and DipN2 the second.

The following table breaks down the participants according to their international status; allowing students from the UK to compared to those who are not.

Degree * UKOrNot Crosstabulation

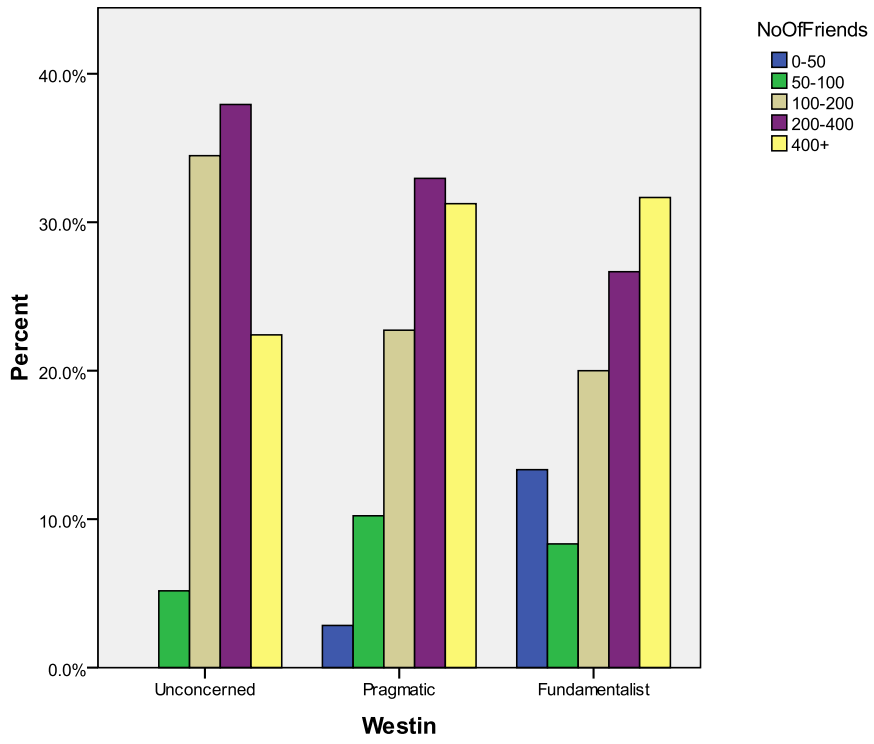
			UKOrNot		
			International	Home	Total
Degree	**Nursing**	Count	10	49	59
		% within Degree	16.9%	83.1%	100.0%
		% of Total	3.0%	14.9%	18.0%
	Spanish	Count	16	18	34
		% within Degree	47.1%	52.9%	100.0%
		% of Total	4.9%	5.5%	10.4%
	Law	Count	17	53	70
		% within Degree	24.3%	75.7%	100.0%
		% of Total	5.2%	16.2%	21.3%
	English	Count	4	51	55
		% within Degree	7.3%	92.7%	100.0%
		% of Total	1.2%	15.5%	16.8%
	DipN	Count	3	41	44
		% within Degree	6.8%	93.2%	100.0%
		% of Total	.9%	12.5%	13.4%
	DipN2	Count	5	61	66
		% within Degree	7.6%	92.4%	100.0%
		% of Total	1.5%	18.6%	20.1%
Total		Count	55	273	328
		% within Degree	16.8%	83.2%	100.0%
		% of Total	16.8%	83.2%	100.0%

The following breaks down the participants according to their age; participants over the age of 21 are listed as mature as per the university classification.

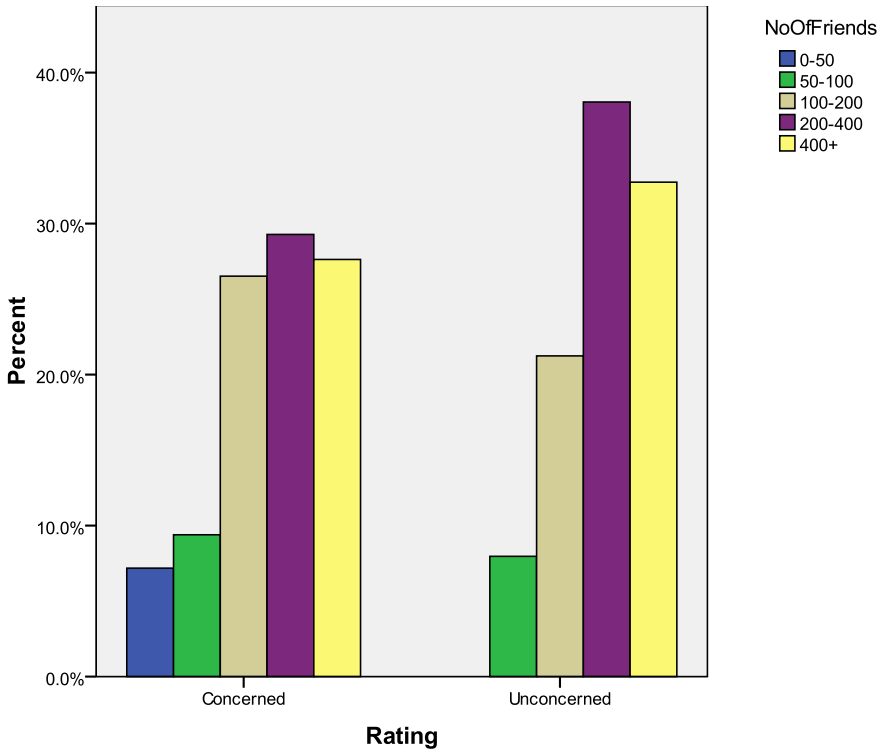
Degree * Mature? Crosstabulation

			Mature?		
			Non	Mature	Total
Degree	**Nursing**	Count	28	33	61
		% within Degree	45.9%	54.1%	100.0%
		% of Total	8.2%	9.7%	17.9%
	Spanish	Count	22	12	34
		% within Degree	64.7%	35.3%	100.0%
		% of Total	6.5%	3.5%	10.0%
	Law	Count	58	12	70
		% within Degree	82.9%	17.1%	100.0%
		% of Total	17.1%	3.5%	20.6%
	English	Count	42	14	56
		% within Degree	75.0%	25.0%	100.0%
		% of Total	12.4%	4.1%	16.5%
	DipN	Count	16	32	48
		% within Degree	33.3%	66.7%	100.0%
		% of Total	4.7%	9.4%	14.1%
	DipN2	Count	35	36	71
		% within Degree	49.3%	50.7%	100.0%
		% of Total	10.3%	10.6%	20.9%
Total		Count	201	139	340
		% within Degree	59.1%	40.9%	100.0%
		% of Total	59.1%	40.9%	100.0%

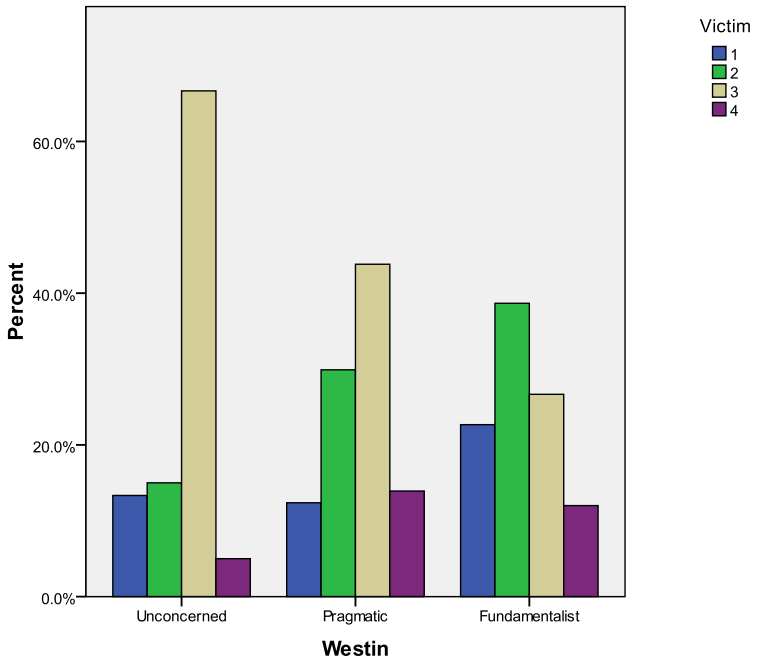
Extra Figures – Concern & Relationships



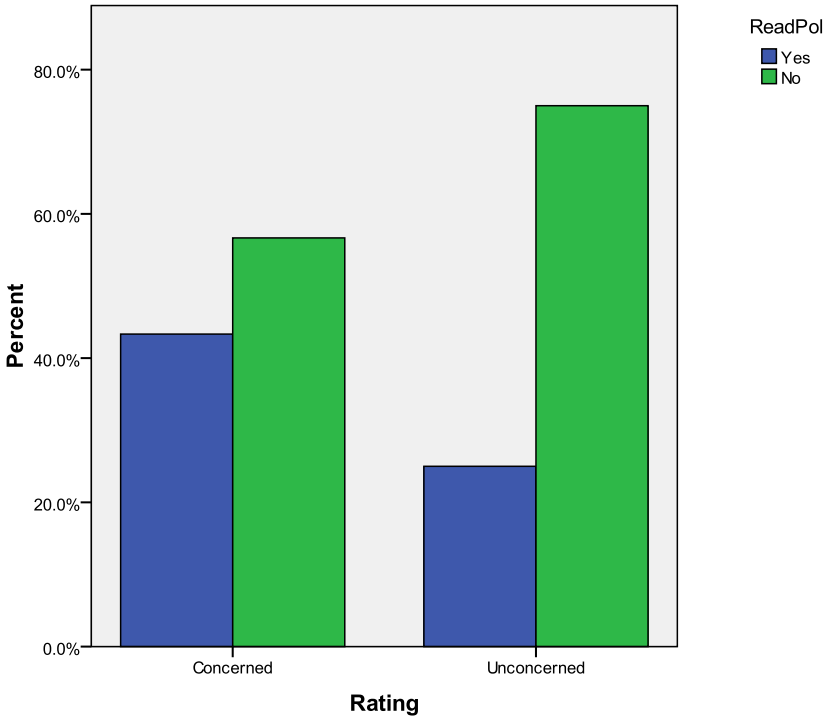
No of Friends across Westin Ratings (chi-square p=.007)



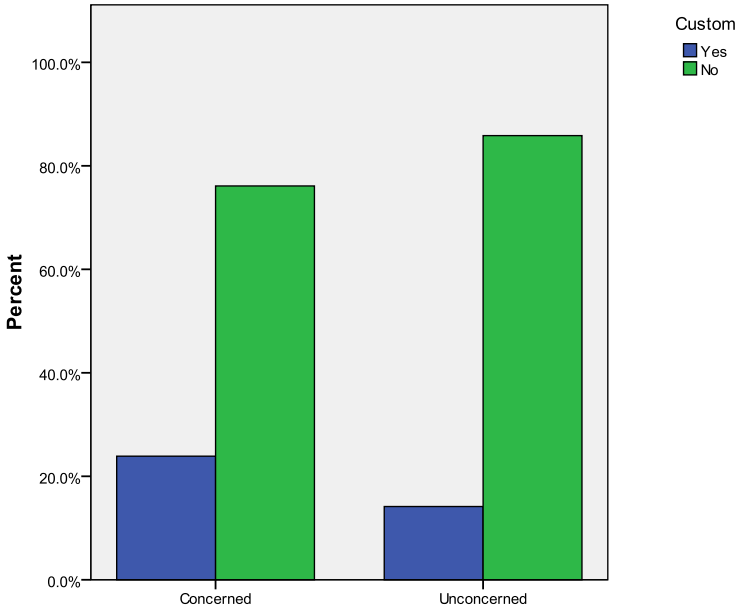
No of Friends across Self-Reported Concern (chi-square p=.023)



Westin Rating and Perceived Victim of Attach (chi-square p<.0001): 1=Yes, 2=Someone I know, 3=Not at all and 4=Do not Know

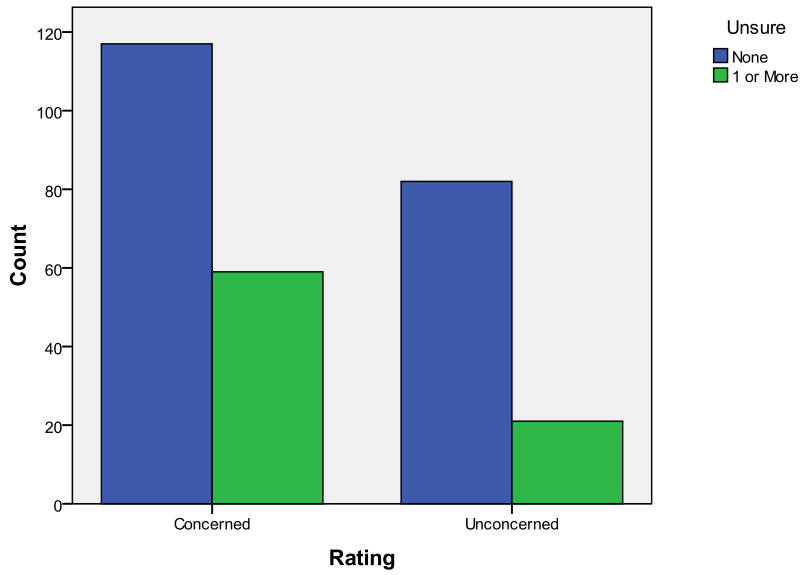


Self-Reported Concern and Policy Engagement (chi-square p=.002)

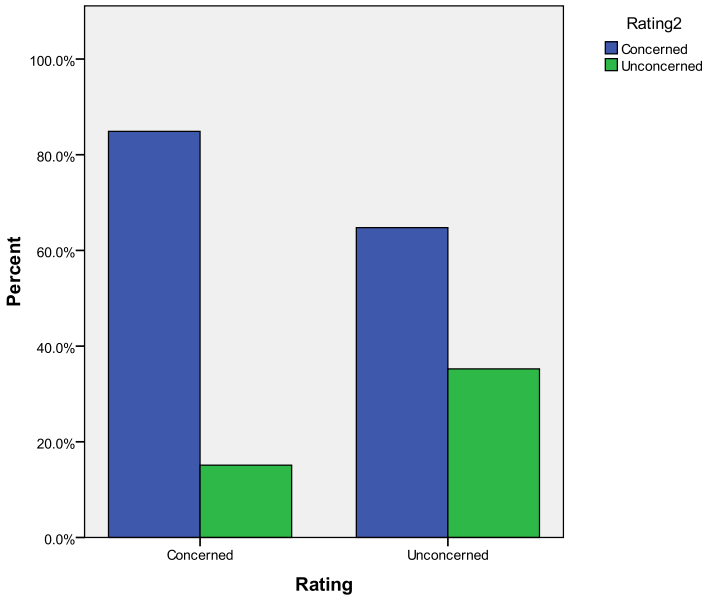


Self-Reported Concern and Reported use of Custom Settings (chi-square p=.043)

Bar Chart



Self-Reported Concern and Reported Unsure of at least 1 setting (chi-square p=.019)



Self-Reported Concern and Concern over Data Ownership (chi-square p<.0001)

Statistics of modal responses to reported settings

		FullName	DoB	Address	Email	Phone	HomeTown	IM	PersonalInfo	RelPol
N	Valid	286	285	267	279	263	281	267	280	278
	Missing	65	66	84	72	88	70	84	71	73
Mode		1	4	4	4	4	4	4	4	4

Second table;

Statistics

		EducationWork	AboutMe	FamilyRelationships	PhotosOfMe	WhoComment	PostsByOthers	PostsByMe	PhotoAlbums
N	Valid	280	280	281	282	282	279	281	281
	Missing	71	71	70	69	69	72	70	70
Mode		4	4	4	4	4	4	4	4

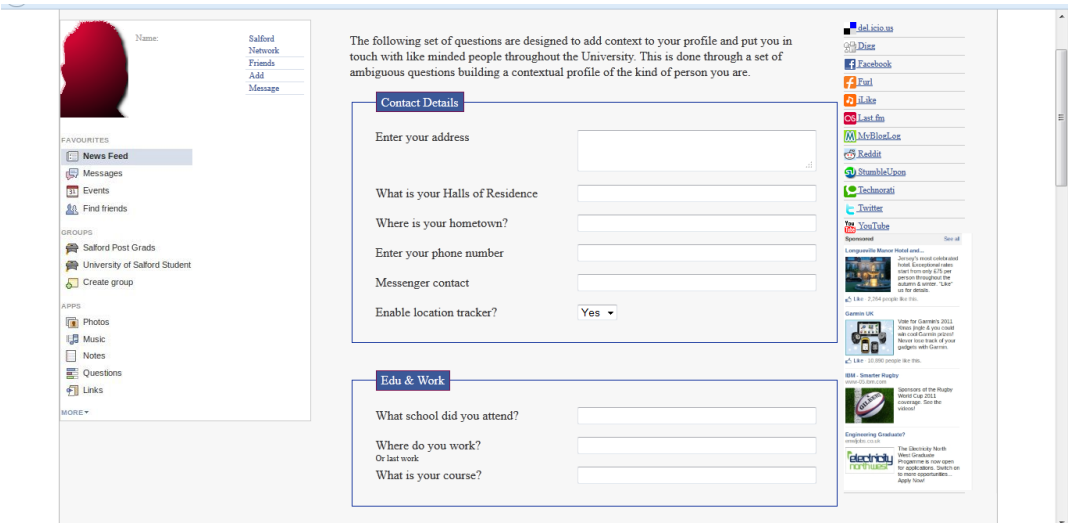
Each of these tables shows that the modal response to the reported settings was 4 in every category (bar full name) which represent “Friends Only” (1 is everyone, 2 is Friends and Networks, 3 is FOAF, 5 is custom and 6 is unsure).

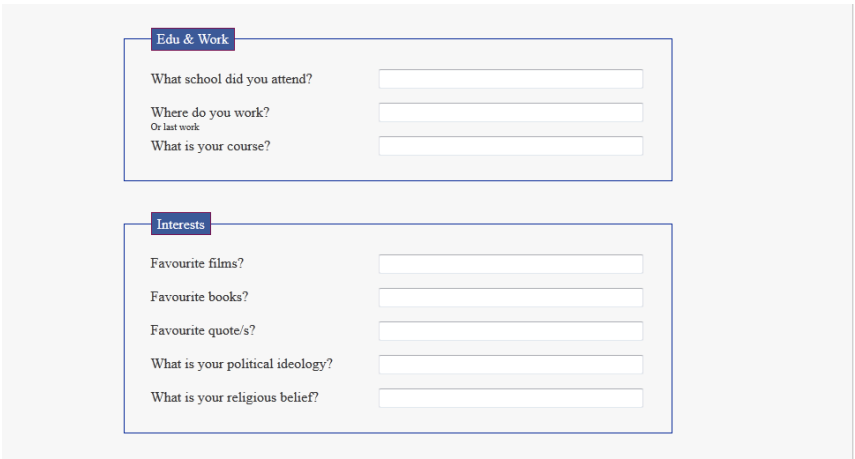
Appendix 4 – Experiment Screenshots

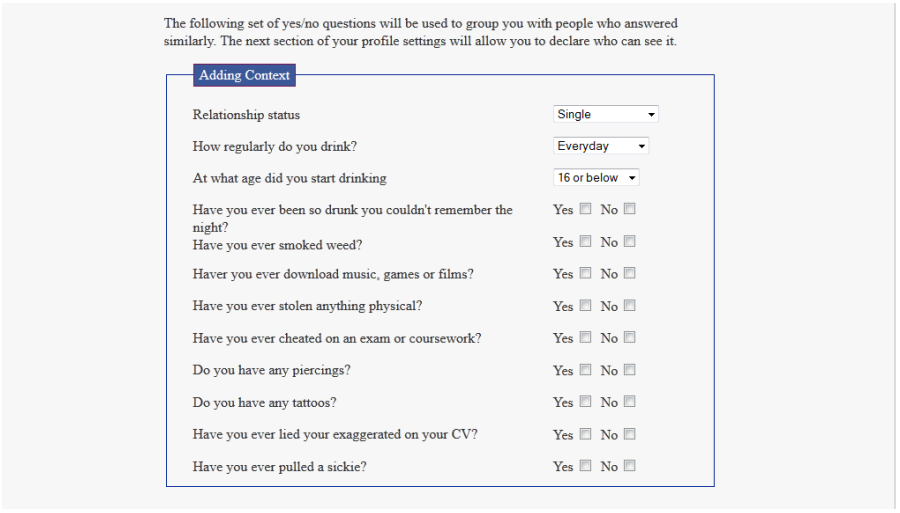
Control Group

First, the initial sign-up screen;

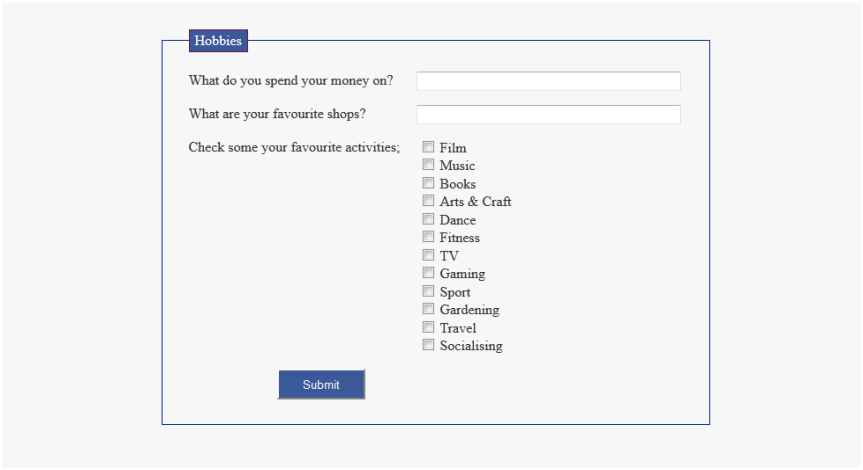
Which leads to the following screen with more personal profile building questions:



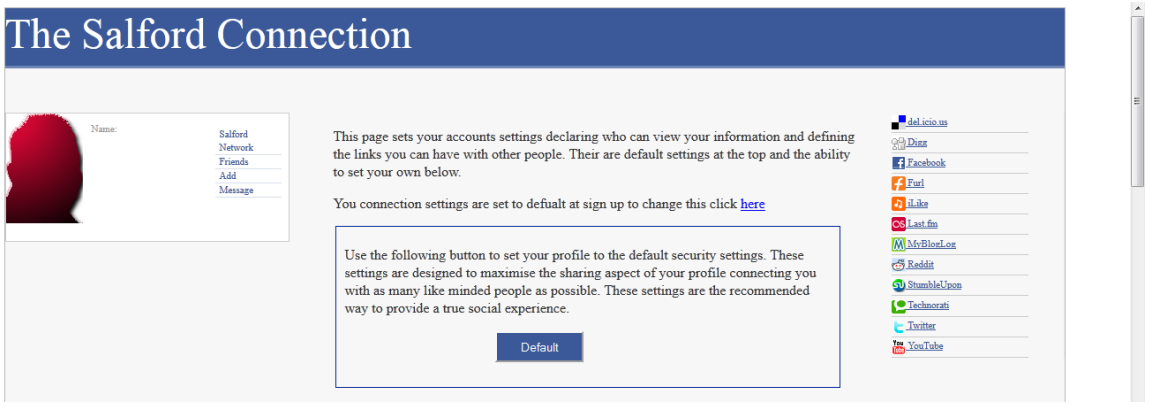




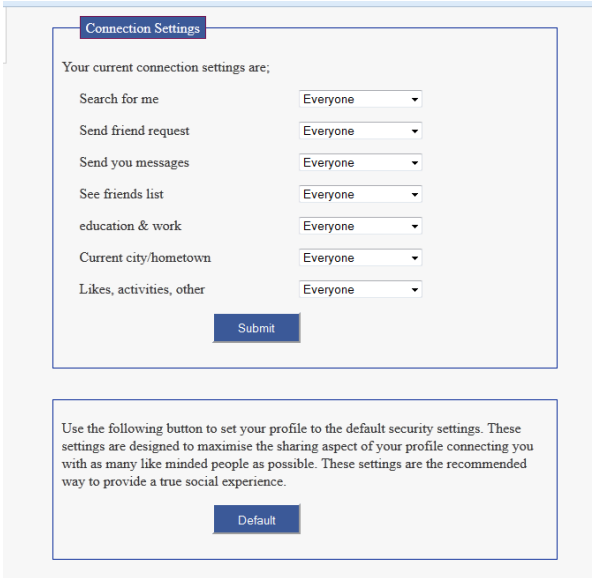
And finally,



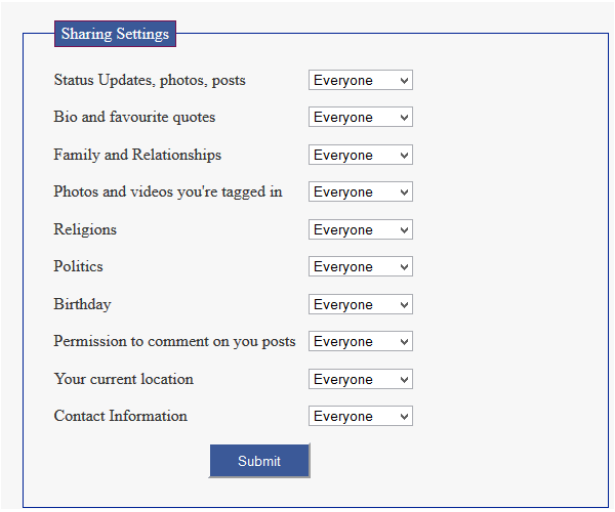
Completion of this page leads to the following page where the participants can set their privacy profile settings;



Note, the first link leading to the setting of the connection portion of the participants profiles (these are automatically set to open by default). These can only be changed by following the link thus measuring the level of engagement with the UI. The second button sets all profile settings to the default which is the most open. These extra settings are as follows:



The page also contains the following for general settings (taken from Facebook);

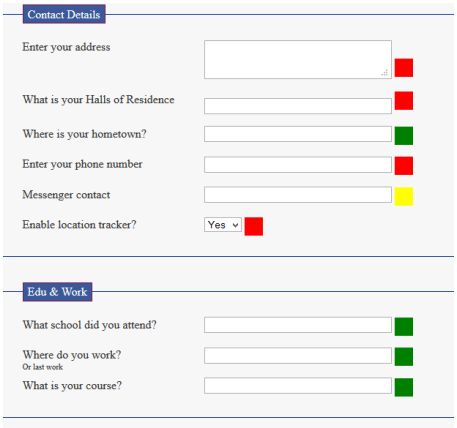


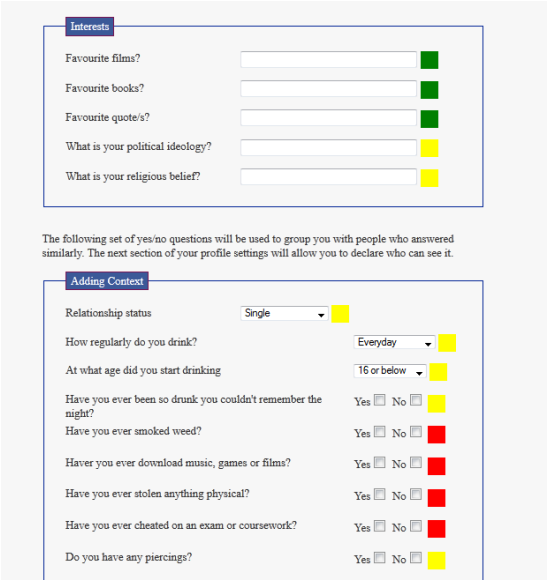
Pressing the submit button completes the sign-up process and finishes the UI experiment portion of the research leading to the exit survey.

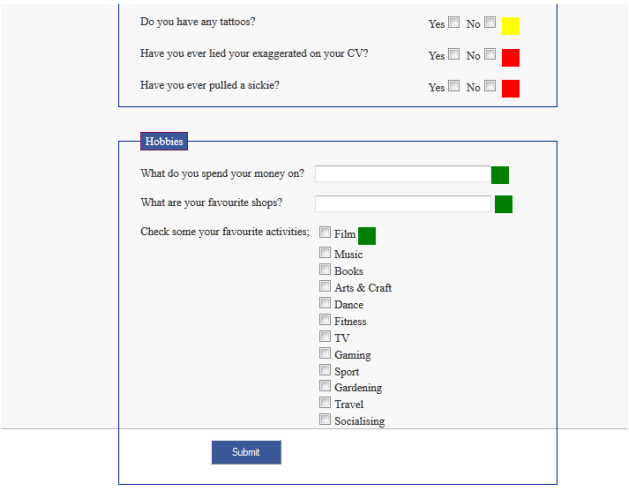
Personal Attitude Group Screenshots

The group follows the exact same structure as above with the following added as the experimental treatments. First, the participants view the following pop-up as an introduction to the system:

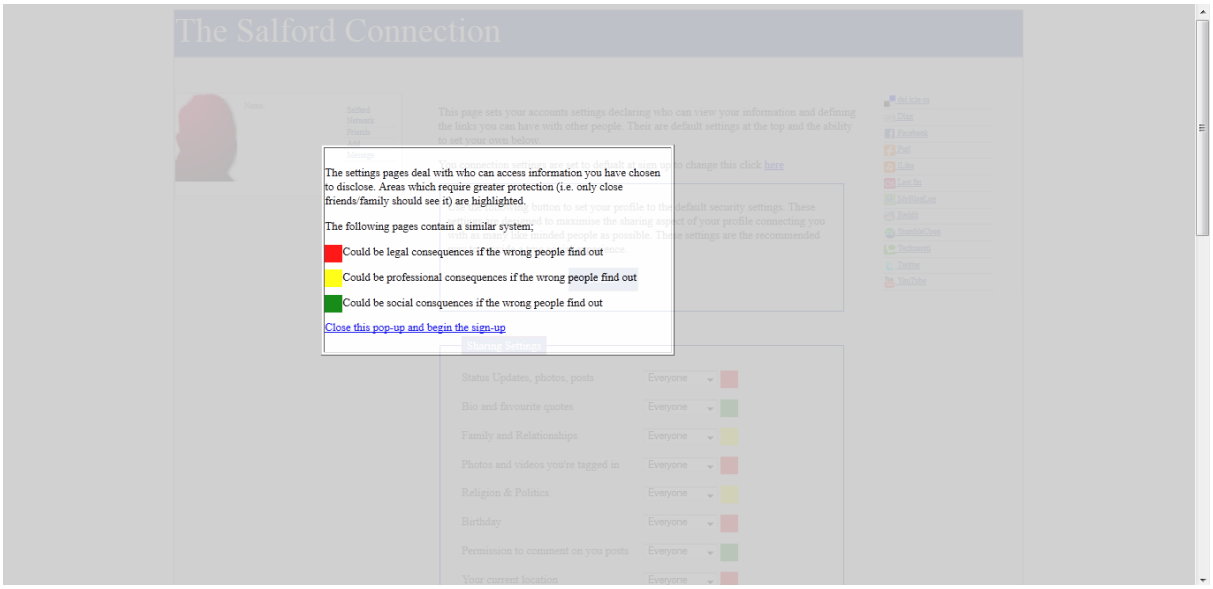
The privacy traffic lights are present as the following indicators of information sensitivity (green for low impact, yellow for medium and red for the highest):



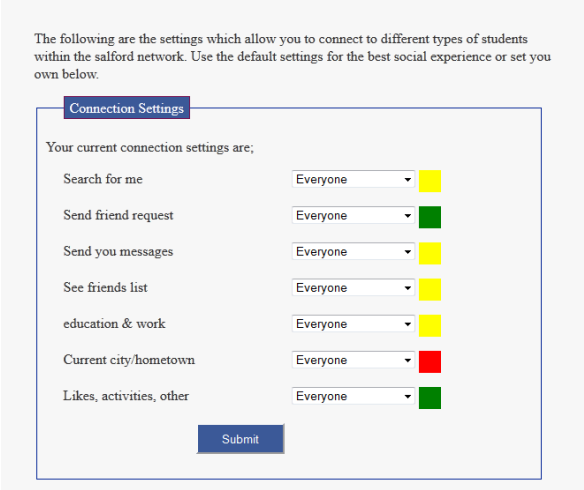




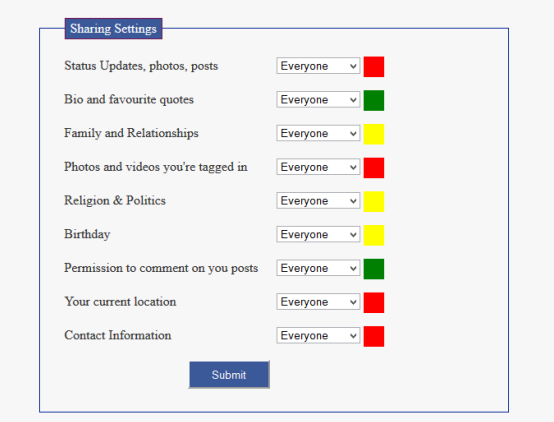
Participants are given a further pop-up for setting their profile privacy:



The connection settings:

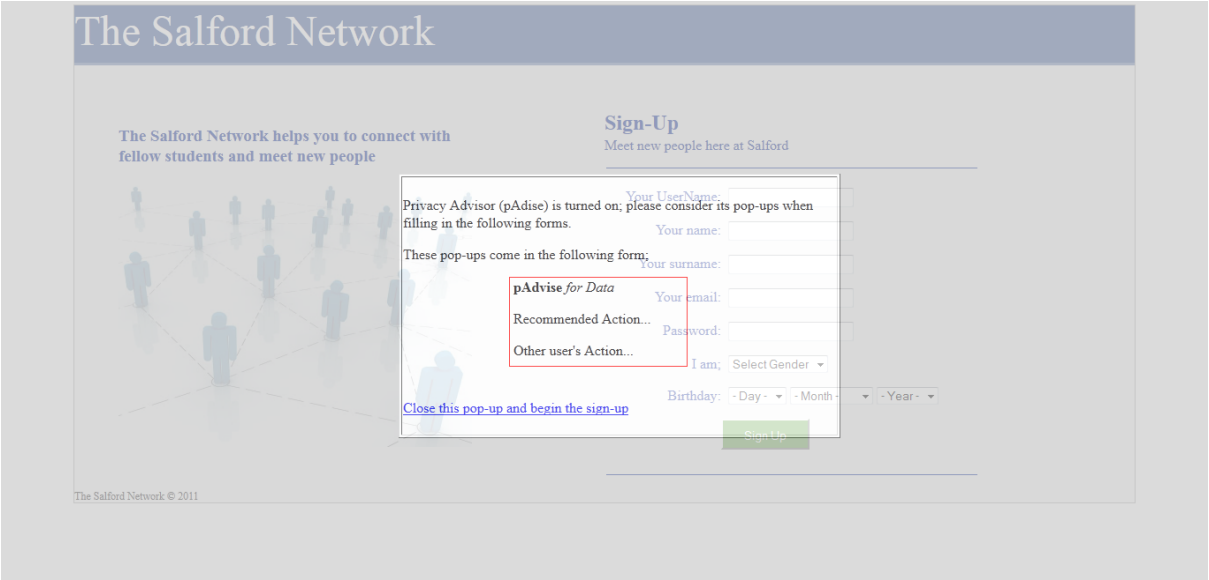


And the profile settings:



The Subjective Norms Group

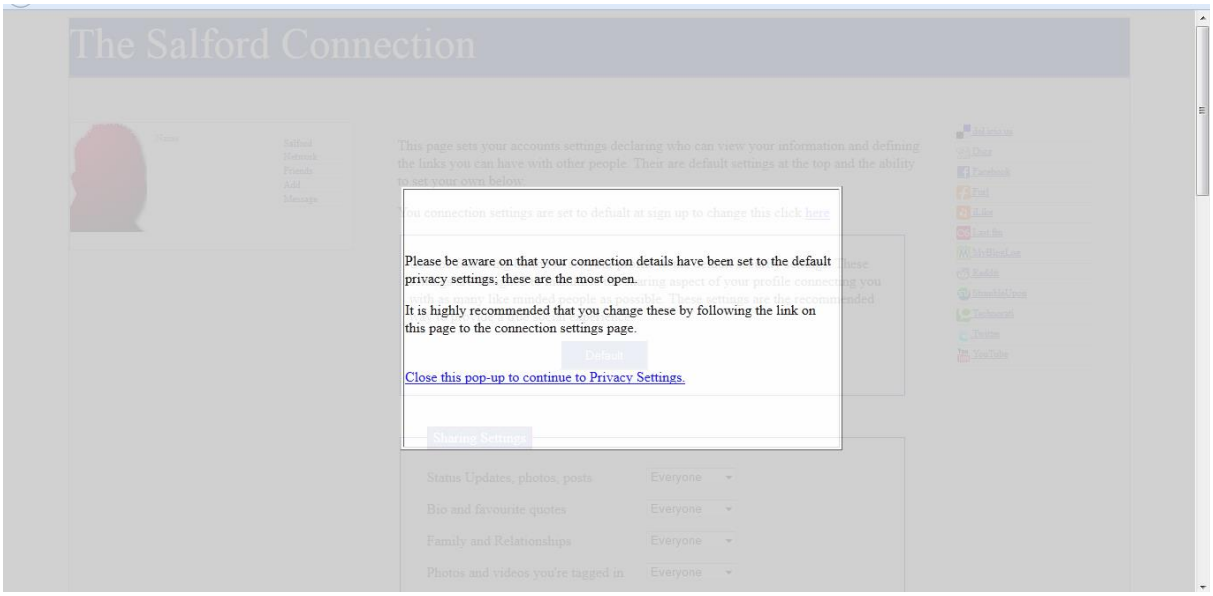
This group contained advice and peer choice incorporated into the selection UI for each questions. Participants are greeted with the following pop-up when starting the experiment;



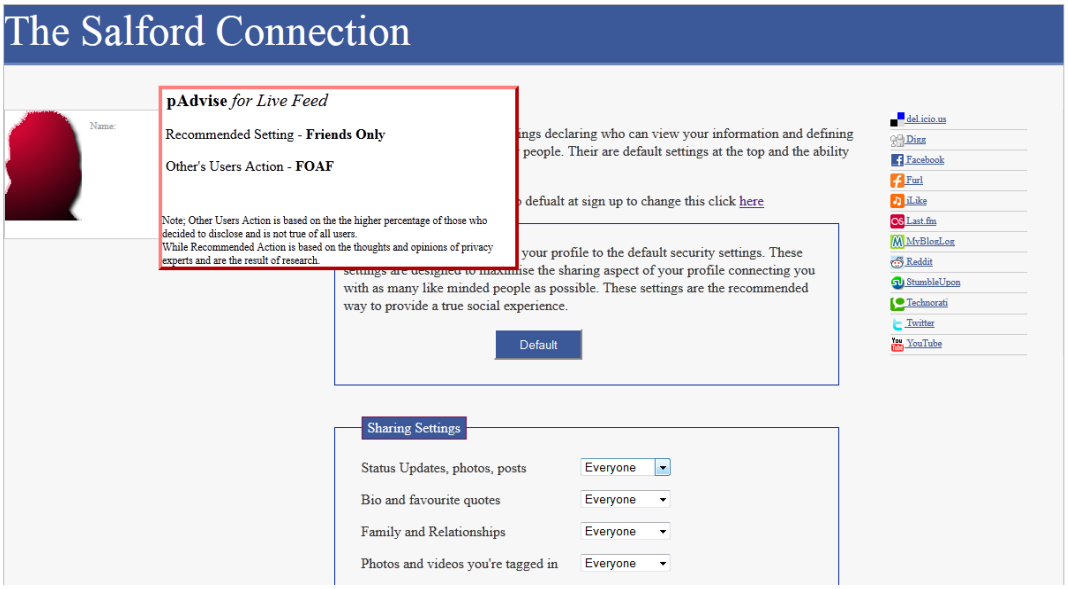
When highlighting a particular data item the following pops up with the two pieces of information for the participant to use in the decision making;

This is present for each data item and the two pieces of information can be found in the complete data item table at the end of this appendix.

At the settings screen, participants are given a similar pop-up informing them of the separate connection settings:

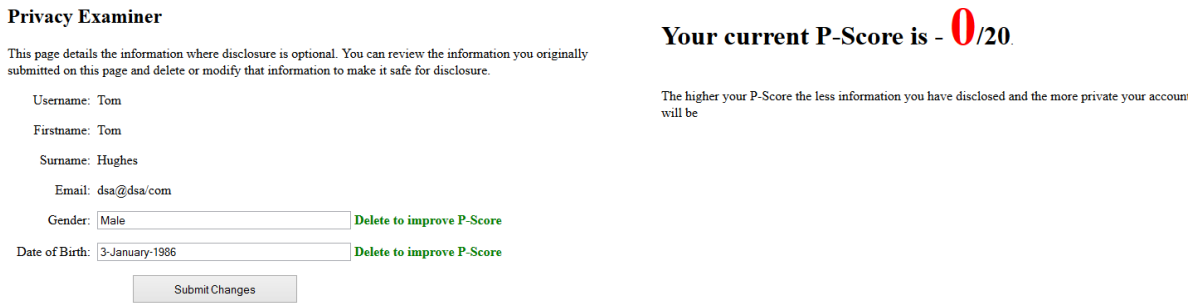


Each setting also has a pop-up advice box (again, full details can be found in the table at the end of the appendix), the advice here gave a recommended setting and the most selected from peer users;

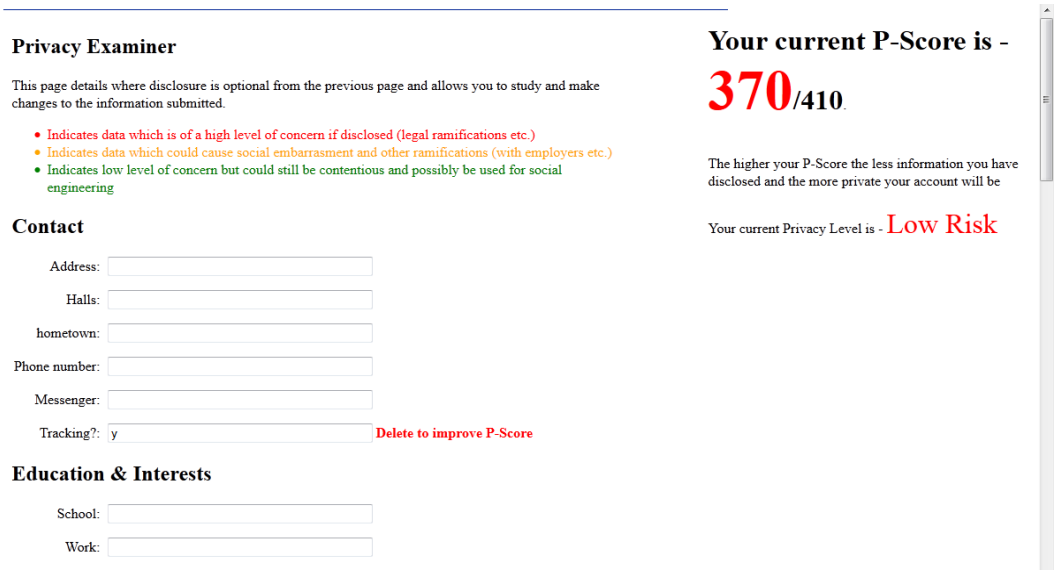


The Perceived Control Group

The Perceived Control group has the same screens as the control group. However, these screens are interspersed with a review screen highlighting the data being submitted by the user. For example, following the sign-up screen the user can review their submissions;

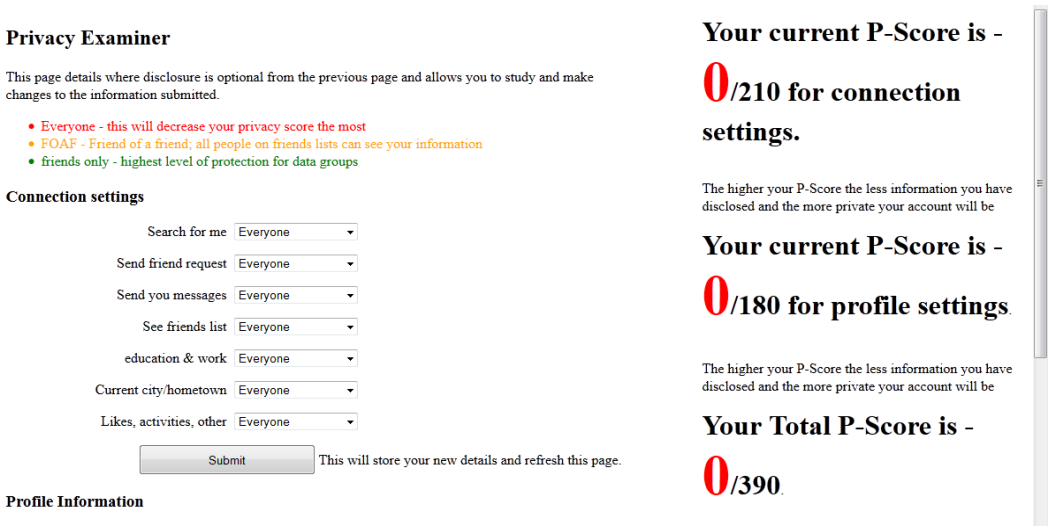


A further example:



Note, the information sensitivity highlighting each piece of data for its level of risk (this is the same as the privacy traffic lights and a full list is found in the table at the end of this appendix).

Finally, the settings:



Note, this is the only experiment where the connection settings are explicitly displayed to the participant; i.e. the settings are shown to the user without their choice to follow the extra link.

Appendix 5 - Information requested

Page 1 – Sign-up			
Data Item	**Sensitivity**	**P-Light**	**P-Light 2**
User Name	N/A	N/A	
First Name	N/A	N/A	
Last Name	N/A	N/A	
Password	N/A	N/A	
Gender	N/A	N/A	
Date of Birth	Medium	Amber	Amber
Page 2 – Profile Builder			
Home Address	High	Red	Red
Term Address	Medium	Red	Red
Hometown	Low	Green	Green
Phone Number	High	Red	Red
Enable Tracking?	High	Red	Red
School	Low	Green	Green
Work Place	Low	Green	Green
Course details	Low	Green	Green
Favourite Films	Low	Green	Green
Favourite Books (TV Shows)	Low	Green	Green
Favourite Quotes (Music)	Low	Green	Green
Are you political, if so which?	Medium	Amber	Amber
Are you religious, if so which?	Medium	Amber	Amber
Relationship Status	Medium	Amber	Amber
Partner Name	Medium	Amber	Amber

Drinking Regularity	Medium	Amber	Amber
Age Started Drinking	Medium	Amber	Amber
Drinking Effects	Medium	Amber	Amber
Drug Use	High	Red	Red
Pirated Media	High	Red	Red
Stolen	High	Red	Red
Cheated	High	Red	Red
Piercings	Medium	Amber	Amber
Tattoos	Medium	Amber	Amber
Lied on a CV	High	Red	Red
Faked Illness	High	Amber	Red
Spending habits	Low	Green	Green
Favourite Shops	Low	Green	N/A
Interests	Low	Green	Green
Personal Email	Medium	N/A	Amber
Sexuality	Medium	N/A	Amber
Donor	Low	N/A	Green
Charity	Low	N/A	Green
Page 3 – Profile Settings			
Status Update, Photos, Posts	High	Red	
Bio and Quotes	Low	Green	
Family and Relationships	Medium	Amber	
Photos and Videos	High	Red	
Religion and Politics	Medium	Amber	
Birthday	Medium	Amber	
Permission to comment	Low	Green	
Current Location	High	Red	

Contact Information	High	Red	
Page 4 – Connection Settings			
Who can search	Medium	Amber	
Who can request friendship	Low	Green	
Who can send messages	Medium	Amber	
Who can see friends list	Medium	Amber	
Education and Work info	Low	Green	
Current Hometown	Low	Green	
Likes and Activities	Low	Green	

Appendix 6 – Consent Form and Instructions

Dear Student,

As part of my PhD research I am conducting experiments using a signup process for social network systems.

I am writing to ask for your participation in a series of experiments revolving around social network use. The social network being tested is a Salford University version of Facebook and shall look and feel very similar. All the information required during the experimentation process is **entirely optional** and shall be kept in secure, encrypted databases**.**

Each group taking part in the experiment will have differences in the system where the actions of the groups involved will be compared for differences; you will be randomly assigned to a group before the experiments begin.

The information stored in the databases from the experiments shall not be viewed by anyone other than researcher and shall be immediately destroyed after the research has been completed. Furthermore, the research shall not analyse the data itself but instead shall look at what data is present and what is absent; therefore, the actual information stored shall not be analysed within a deeper context than that. **You will have the opportunity to set who can see what information or not divulge any information you do not wish to.** It is important to note that you will be given a participant number which will be used to relate to the experiments rather than your name; this is to assure anonymity from the data involved.

At any stage of the research **you have the rights to withdraw your information.** Throughout the study **you also hold the complete rights to ask any further questions regarding the study which occur to you.**

When the study is concluded **you will be given access to a summary of the findings of the study and these shall be made available on request where publication has not been possible.**

If this is agreeable to you, please sign the attached research consent form.

I would like to take this opportunity to thank you for taking the time to help me with my research. If you have any queries, please feel free to contact me at the following email;
T.Hughes-Roberts@edu.salford.ac.uk

Yours Sincerely,

Thomas Hughes-Roberts
PhD Student
The Computer Networking and Telecommunications Research Group
School of Computing, Science and Engineering
University of Salford

User Behaviour in Online Social Networks; an Examination of User Action during the Sign-up Process

Principal Researcher: Thomas Hughes-Roberts

School of Computer Science and Engineering

1. I confirm that I have read and understand the information provided for the above study. I have had the opportunity to consider the information, ask questions and have these answered satisfactorily. ☐

2. I understand that my participation is entirely voluntary and that I am free to withdraw at any time, without giving reason and that this will not affect my legal rights. ☐

3. I understand that any personal information collected during the study shall be kept entirely confidential. ☐

4. I agree to take part in the study. ☐

Researcher Name: Thomas Hughes-Roberts **Signature:**	**Date:**
Participant Print Name: **Signature:**	**Date:**

Instructions for Experiments

You have been asked to join a brand new social network aimed exclusively at NTU students. You will be following the account creation process for your new account which creates your specific network based on how much information about yourself you supply. To begin follow the instructions below and sign the back of this form to agree to your participation.

1. Open a web browser.

2. Enter the following in the address bar of the browser; *Variable address*

3. Follow the sign-up process and create your account

Appendix 7 – Experiment One Results

Group	Gender (M/F)
Control	8/2
Personal Attitude	8/3
Perceived Control	11/1
Subjective Norms	7/5

The following tables indicate the amount of questions left out in each category asked of participants during the experiment, giving a total number for each participant (provided in the "Total" column). The total amount of settings is also provided out of a maximum of 200.

Control								
Users	**Signup**	**Contact**	**Eduinterest**	**Context**	**Marketing**	**Total**	**Settings**	**Connection**
1	0	0	6	0	0	6	200	0
2	0	0	0	0	0	0	80	0
3	0	0	4	0	0	4	0	0
4	0	1	0	0	0	1	200	0
5	0	2	0	0	0	2	0	0
6	0	2	5	1	2	10	0	0
7	0	1	0	0	0	1	200	0
8	0	3	6	1	2	12	200	0
9	0	0	0	0	0	0	0	0
10	0	1	0	0	0	1	200	0
					Total	37		0
					Average	3.7		
					St.Dev	4.295993		

Attitude

Users	Signup	Contact	Eduinterest	Context	Marketing	Total	Settings	Connection
1	0	5	6	1	0	12	150	0
2	0	5	6	0	0	11	200	0
3	0	0	0	0	0	0	40	0
4	0	4	8	0	3	15	0	0
5	0	5	8	9	3	25	200	0
6	0	0	4	0	0	4	110	0
7	0	4	8	0	1	13	200	0
8	0	0	4	1	0	5	200	0
9	0	4	2	3	0	9	100	0
10	0	1	4	1	2	8	200	0
11	0	4	1	0	0	5	200	0
						Total	107	
						Average	9.727273	
						St.Dev	6.739301	

P-Control1

Users	Signup	Contact	Eduinterest	Context	Marketing	Total	Settings	Connection
1	0	1	3	1	0	5	200	0
2	0	3	5	4	2	14	200	0
3	0	2	2	0	0	4	140	0
4	0	1	8	2	0	11	200	0
5	0	0	8	0	0	8	200	0
6	0	4	8	9	3	24	0	0
7	0	4	8	9	3	24	0	0
8	0	4	2	0	0	6	200	0
9	0	4	3	0	1	8	180	0
10	0	4	4	0	0	8	200	0
11	0	4	6	9	3	22	200	0
12	0	4	4	3	1	12	200	160
						Total	146	
						Average	12.16667	
						St.Dev	7.321616	

P-Control2								
Users	**Signup**	**Contact**	**Eduinterest**	**Context**	**Marketing**	**Total**	**Settings**	**Connection**
1	0	1	3	0	0	4	200	0
2	2	4	6	12	2	26	200	170
3	2	5	8	11	3	29	80	60
4	2	5	8	12	3	30	200	210
5	2	1	8	0	0	11	200	0
6	2	5	8	12	3	30	200	210
7	0	4	8	9	3	24	0	0
8	2	5	2	0	0	9	200	0
9	0	4	3	1	1	9	180	150
10	0	5	4	0	0	9	200	0
11	0	4	6	9	3	22	200	150
12	0	4	5	10	1	20	200	160
						Total	223	
						Average	18.58333	
						St.Dev	9.596006	

Subjective								
Users	**Signup**	**Contact**	**Eduinterest**	**Context**	**Marketing**	**Total**	**Settings**	**Connection**
1	0	1	1	0	0	2	200	0
2	0	3	1	0	1	5	0	0
3	0	1	0	0	0	1	0	0
4	0	4	5	10	3	22	200	120
5	0	3	0	3	0	6	0	0
6	0	4	5	0	0	9	0	0
7	0	1	4	0	0	5	0	0
8	1	4	6	11	1	23	200	0
9	0	3	6	0	1	10	190	0
10	0	2	6	0	1	9	0	0
11	0	5	7	6	2	20	0	0
12	0	3	7	0	0	10	80	0
						Total	122	
						Average	10.16667	
						St.Dev	7.541803	

Appendix 8 – Interview Notes

Potential Interview Questions

The following are a set of individual questions which could be asked following participation in the experiments; note, that these are informal and deviation from the questions are encouraged should an interesting observation be made and pursued.

1. What did you think of the system?
2. Did you disclose all the fields asked of you?
3. If so, why?
4. If not, why?
5. What did you think of the changes to the User Interface?

The following are notes taken during the observations and interviews.

Control Group Notes from Observations and Interviews

"I have been honest" – participants view it as a competition to see who could be the most risqué about themselves, disclosing very sensitive information in droves.

"I gave everything because the system asked me to" – participants believed that they had to disclose all the information asked of them in order to appease the system and complete the goal of signing up.

"I thought information was optional as there was no red asterisk next to the field" – despite this the participant gave a lot of information.

"I hope the police don't see this" – aware that they are disclosing information about themselves that is extremely sensitive and yet are still seemingly happy to do so.

In terms of the privacy settings applied; participants mainly (except for two instances) applied friends only protection. Around 50% of participants applied no protection at all and seemed to be "clicking through at this point as they looked to see that some settings had already been applied by default and chose to leave it in order to finish the task at hand. Furthermore, participants who did select settings seemed to do so quickly without apparent awareness of what they were applying that setting to. "Friends only" was selected with practiced ease and obvious familiarity from those that did apply settings. This would suggest that participants feel they know what should be selected and did so habitually and without obvious thought in a new environment. HCI metaphors and self-efficacy could both be used to explain the two described observations here.

Participants were also very focussed on each specific question, scrolling across what asked of them with the cursor; suggesting that each question was taken granularly and not seen in the wider context of what has been asked of them i.e. they seemed more concerned with getting the answer right than deciding if they should answer.

Personal Attitude Group Notes

Initial results from some interviews – Did you agree with the lights? Generally, yes; "made you realise that there was the option of disclosing information and helped to choose what to disclose".

Participants had the most concern about directly relatable information (Address, phone no. etc.), least concern regarding contextual questions – "lies on CV is ok, everyone does it, so it's ok to tick the box and disclose it". Through further probing the participants also said; "People can just lie about these questions so I see no reason not to answer them". This came about when it was explained that, even if lied about, there is still information present which can be used to form an opinion about the user so why not just leave it black. This did not seem to be an option when in system was the response from the participant.

However each participant felt that the lights were beneficial overall, providing the prompt reminding the participants of choice.

Perceived Control Notes

Participants were heard saying the following things - "does this mean the data is optional then?" "Should I give it?" and "I'll delete this to get my P-Score down".

This shows that participants were thinking more than other groups about the pieces of data within a privacy context, actually asking themselves the question of whether or not they should disclose it.

One participant after being interviewed felt that the salient feature made very little difference to how they behaved within the network (this was participant #12). This is despite a change in the total number of disclose field to 20 from 10. So his mind was clearly changed through the course of the process.

In the after interview one participant stated that he felt "if it's easy to disclose I will" which gives some indication of how Facebook might persuade people to disclose information.

Subjective Norms Notes

When asked why they gave some pieces of information a participant responded – "I gave it but wish I didn't, I just answered the easy ones, they required less time/effort". This is response specifically to the response given to the contextual questions which were mainly tick boxes. When asked why admitted to the various activities these question dealt with the above was used to explain their actions. Again this shows how the UI can be persuasive in getting the user to disclose information.

Participants seemed to be put off by the initial screen lock and thought there was an error in the way in which they were using the system. One participant was heard exclaiming; "what have done?". The behaviour of "clicking through" was also clear here with participants simply filling in and clicking in order to get to the next screen. Did the message affect their efficacy?

As seen in the control group, participants were very engaged with each question individually, scrolling over it with the cursor and ensuring they filled it in correctly. The problem seemed to be that when the initial popup advisor appeared it caused some concern but was quickly ignored as an affectation of the system. Through the focus on each question the information contained in the pop-up was not taken on board. Participants literally got closer to the screen and tackled the question on its own completely isolated from the rest of the UI. The fact that the information appears away from the information request did not help; perhaps it should appear next to it like in personal attitude which had a greater effect (perhaps due to its simplicity and conceptualisation of a traffic light metaphor).

A further point can be made for conceptualising the system which demonstrates the importance of self-efficacy; a participant said “I wasn’t sure how to handle it until I just thought of it as Facebook” This was in response to being asked how they decided what to disclose to the system.

Appendix 9 – Participant Information, Westin Tables and Exit Survey Data for Experiment Two

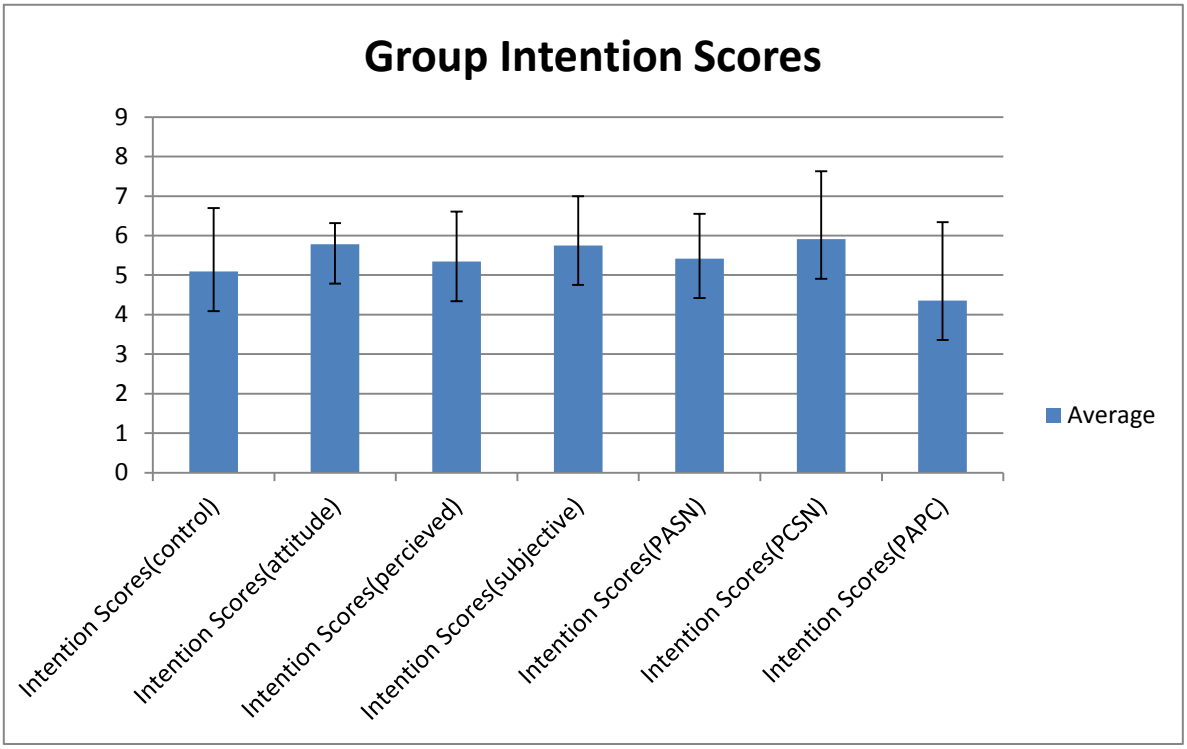
Group	Gender (M/F)
Control	16/4
Personal Attitude	17/6
Perceived Control	17/4
Subjective Norms	19/4
PA+SN	11/3
SN+PC	13/1
PA+PC	12/2

Spread of Westin Ratings Across Experiment 2 Groups

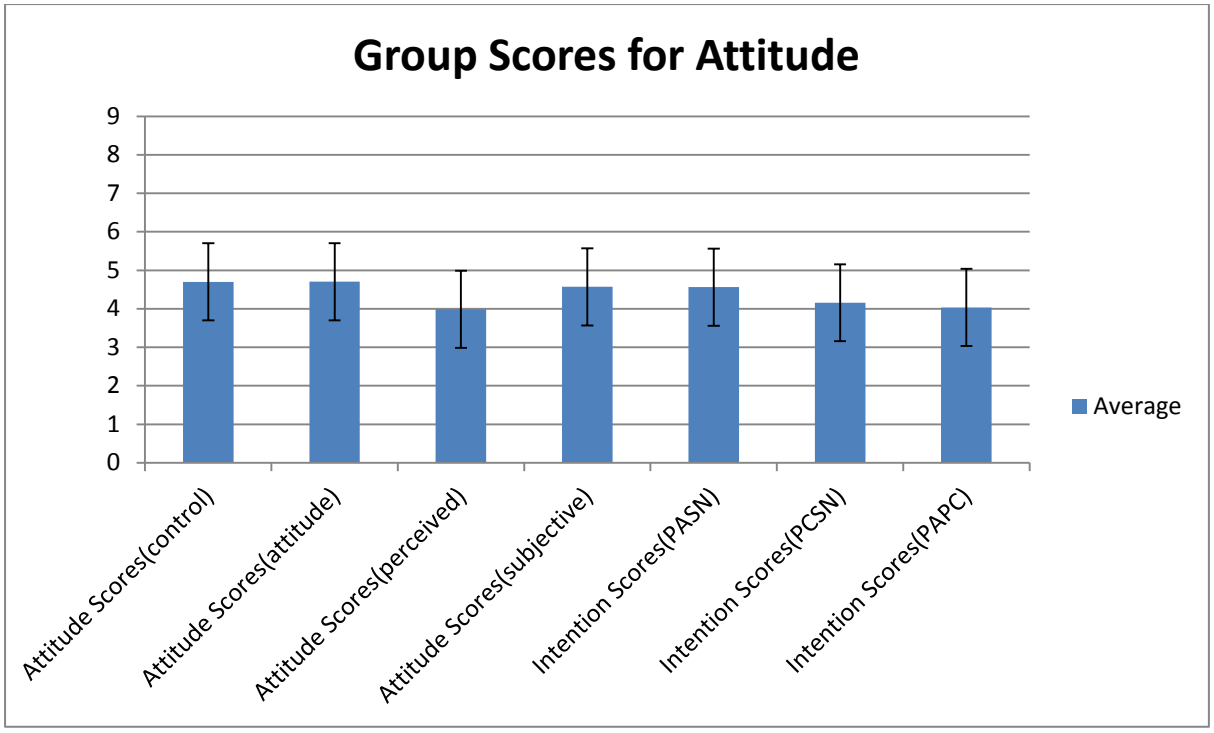
Group	Fundamentalist (number/%)	Pragmatist (number/%)	Unconcerned (number/%)
Control	12.5%	87.5%	0%
PA	18.75%	75%	6.25%
SN	33.3%	61.1%	5.6%
PC	26.3%	73.7%	0%
PA+SN	41.7%	50%	8.3%
SN+PC	36.4%	54.4%	9.1%
PA+PC	42.9%	28.6%	28.6%
Wider Research	25%	57%	18%

It is interesting that the treatment groups held far higher privacy fundamentalists than the control group which mainly held pragmatist rated participants. This could be due to participants being influence by the treatments present in the experiment and their level of privacy being raised. However, it could also be due to variation among the groups and it is not clear from this measure alone what the effect of the treatment is on participant's privacy perception.

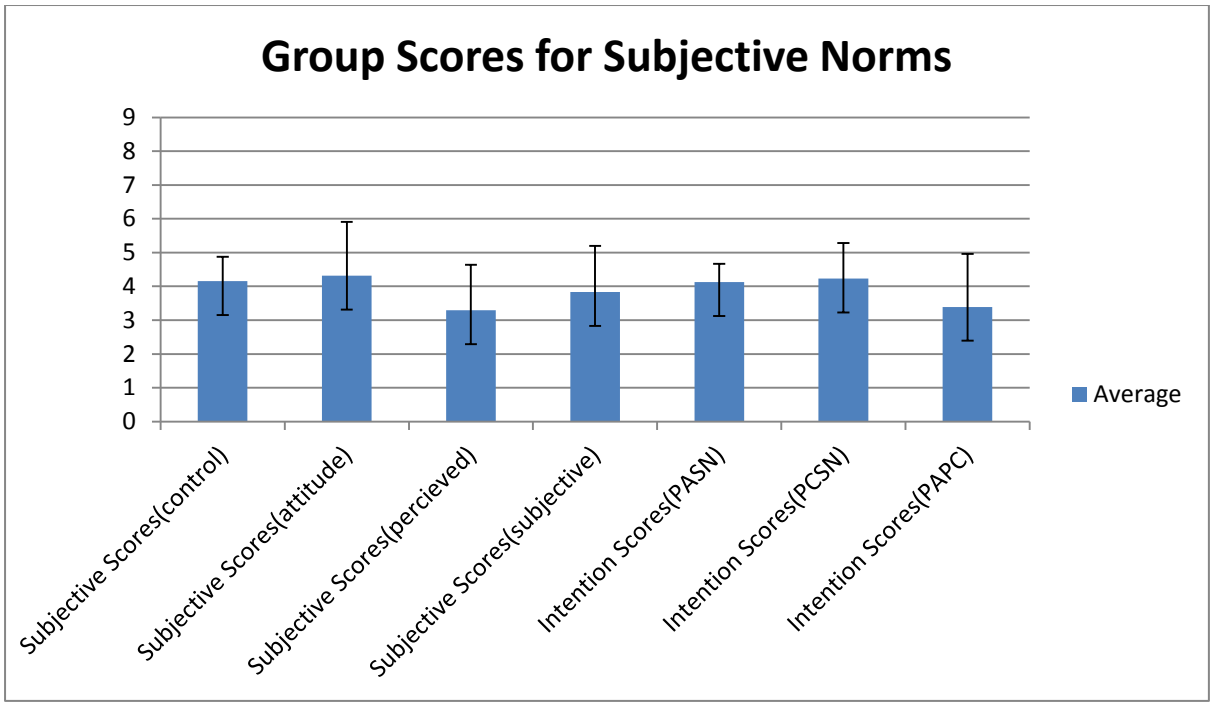
The following summarises the TPB measures of intention, attitude, subjective norms and control for the experiment two groups.



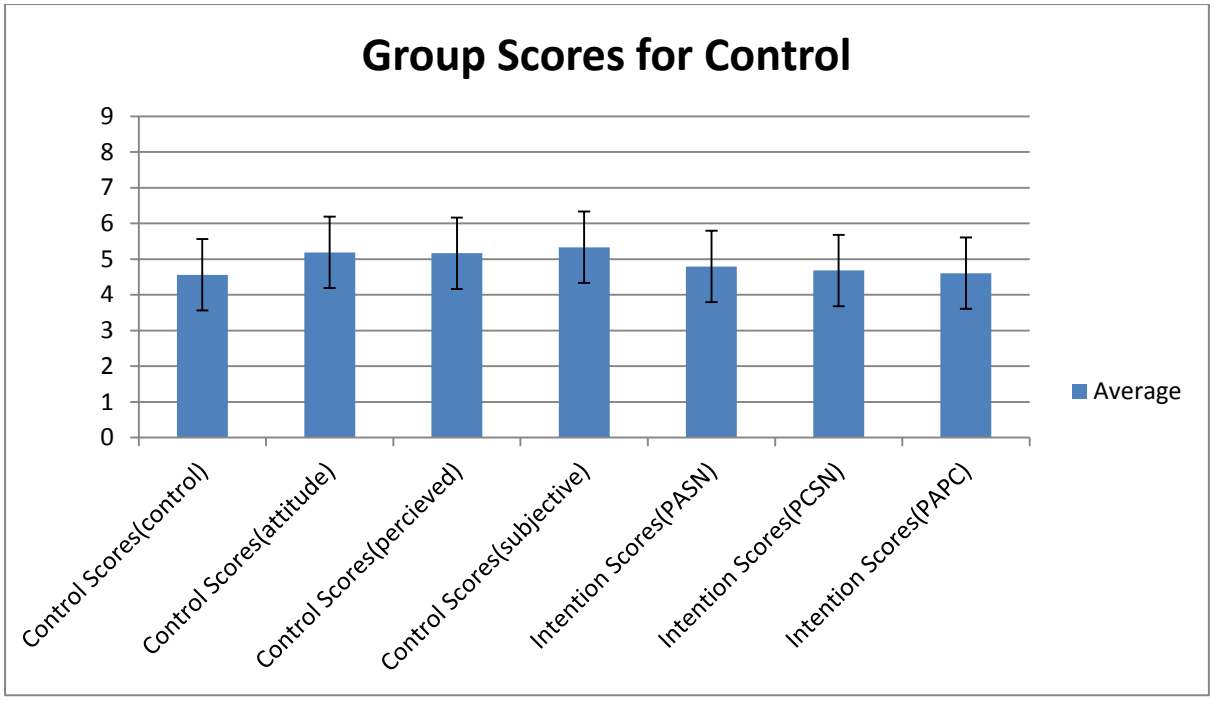
As before a higher score would suggest that participants wish to disclose as little as possible and it is interesting the PAPC group which disclosed the least in the experiment declared held the lowest intention score. Again, this may be due to variation as there is no statistically significant difference compared to the control (p=.536). As such, there is further evidence here that a generic survey approach is inefficient in exploring privacy intention.



A larger attitude score would suggest that the there is a positive attitude toward disclosure. There is very little difference between the groups; although, the PC groups are the lowest which also had the lowest levels of disclosure when compared to the control.



It is interesting that the groups which held the lowest levels of disclosure have the lowest attitude scores (PC and PAPC) (which suggest an unwillingness to listen to advice). However, these differences are not statistically significant.



A higher score would indicate that it was easy to identify and protect sensitive information. Notice the similarity between the single factor groups where disclosure was lessened compared to the control. The factorial groups, however, are broadly similar to the control despite also having less disclosure as in the single factor treatments. These elements of the survey then suggest a need for a more granular prediction if a survey measure is used. However, it may very difficult to create a context that mirrors reality in order to specific behavioural predications regarding privacy behaviour and hence a survey method may not be the most efficient way of exploring the privacy phenomena.

Appendix 10 – Experiment Data Overview

Control

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	0	0	0	0	0	0	160
2	0	2	2	1	0	5	0	0
3	0	1	0	0	0	1	0	0
4	0	1	1	0	0	2	0	120
5	0	0	0	0	0	0	0	60
6	0	4	4	1	1	10	0	260
7	0	1	2	0	0	3	0	260
8	0	1	4	0	0	5	0	0
9	0	6	7	12	4	29	120	260
10	0	0	0	0	0	0	0	0
11	0	2	7	1	0	10	0	260
12	0	5	6	5	0	16	0	0
13	0	0	0	1	0	1	0	100
14	0	1	5	1	1	8	0	0
15	0	0	1	1	2	4	0	40
16	0	0	1	0	0	1	0	240
17	0	0	0	0	0	0	0	260
18	0	1	3	0	0	4	0	260
19	0	5	0	1	4	10	0	260
20	0	5	0	1	0	6	0	260
					Group Average	5.75		
					St.Dev	7.010331		

Personal Attitude

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	3	2	7	0	12	60	180
2	0	1	1	2	0	4	0	260
3	0	2	2	7	0	11	0	0
4	0	3	3	1	0	7	0	260
5	0	1	6	10	1	18	0	260
6	0	4	5	10	1	20	0	260
7	0	2	5	2	1	10	0	0
8	0	3	4	2	0	9	0	0
9	0	4	4	5	1	14	0	0
10	0	3	6	3	4	16	0	80
11	0	4	3	8	0	15	0	260
12	0	3	4	7	0	14	0	110
13	0	3	2	9	0	14	0	120
14	0	2	0	1	0	3	0	120
15	0	4	1	1	1	7	0	40
16	0	3	2	11	0	16	0	260
17	0	2	2	2	0	6	0	0
18	0	5	5	1	4	15	0	180
19	0	5	6	1	0	12	0	160
20	0	3	2	1	0	6	0	220
21	0	5	2	1	3	11	0	260
22	0	4	1	9	0	14	0	260
23	0	3	2	1	0	6	0	80
					Group Average	11.30435		
					St.Dev	4.63617		

Subjective Norms

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	5	2	9	0	16	0	260
2	0	5	2	2	0	9	0	50
3	0	5	2	10	4	21	0	260
4	0	5	8	10	4	27	0	0
5	0	1	3	3	1	8	0	260
6	1	5	3	3	0	12	0	90
7	1	5	8	12	1	27	0	260
8	0	5	2	9	1	17	0	260
9	0	4	2	9	4	19	0	0
10	0	2	2	11	0	15	0	140
11	1	5	2	10	1	19	0	190
12	0	6	8	2	0	16		
13	0	2	5	1	0	8	0	260
14	0	2	1	1	0	4	0	0
15	0	2	0	11	0	13	0	90
16	0	2	2	0	1	5	0	260
17	0	2	2	9	0	13	0	10
18	0	4	2	13	0	19	0	0
19	0	2	2	0	0	4	0	70
20	0	5	1	1	0	7	80	260
21	0	6	2	10	1	19	0	260
					Group Average	14.19048		
					St.Dev	6.845576		

Perceived Control 1

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	0	0	0	0	0	0	120
2	0	1	1	0	0	2	0	180
3	0	1	0	0	0	1	0	0
4	0	6	8	10	4	28	0	260
5	0	1	2	0	0	3	0	0
6	0	0	3	2	0	5	0	260
7	0	0	2	12	0	14	0	140
8	0	0	0	1	0	1	0	80
9	0	1	4	1	0	6	0	140
10	0	0	4	8	0	12	0	260
11	0	3	1	0	0	4	0	40
12	0	2	5	1	1	9	0	0
13	0	3	8	13	4	28	0	260
14	0	1	4	10	0	15	0	260
15	0	1	6	1	0	8	0	140
16	0	1	2	0	0	3	0	160
17	0	1	6	10	4	21	0	150
18	0	0	0	1	1	2	0	130
19	0	1	0	2	0	3	0	260
20	0	5	0	1	1	7	0	260
21	0	1	0	1	1	3	0	260
					Group Average	8.333333		
					St.Dev	8.463648		

Perceived Control 2

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	5	2	13	0	21	120	240
2	0	2	3	10	0	16	120	260
3	1	6	8	14	4	33	120	260
4	0	6	8	13	4	31	0	260
5	0	1	3	0	0	6	0	0
6	0	0	4	2	0	7	0	260
7	0	1	2	13	0	16	80	120
8	0	2	2	11	0	15	0	80
9	0	5	6	8	1	20	0	150
10	0	0	6	8	0	14	30	260
11	0	4	3	0	0	8	0	40
12	0	3	7	10	1	21	0	0
13	1	6	8	13	4	32	120	260
14	0	4	4	13	0	21	90	260
15	0	1	6	1	0	8	0	140
16	1	5	3	13	0	22	100	160
17	1	6	8	13	4	32	0	150
18	1	6	0	1	1	9	0	130
19	0	1	0	0	0	1	120	260
20	1	5	0	2	1	9	60	260
21	1	5	5	3	1	15	0	260
					Group Average	16.71429		
					St.Dev	9.623632		

PASN

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	3	2	9	0	14	40	170
2	0	4	1	0	0	5	0	130
3	0	6	8	9	4	27	0	0
4	0	3	5	0	4	12	0	110
5	0	2	6	10	4	22	0	150
6	0	4	6	9	4	23	0	120
7	0	1	1	1	0	3	0	260
8	0	5	8	10	4	27	0	160
9	0	0	3	1	0	4	0	0
10	0	1	0	3	0	4	0	160
11	0	1	2	2	0	5	0	180
12	0	5	2	11	0	18	100	240
13	0	4	2	9	0	15	0	260
14	0	4	3	10	0	17	0	240
					Group Average	14		
					St.Dev	8.753021		

PAPC1

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	3	2	8	0	13	0	130
2	0	5	2	9	0	16	120	260
3	0	4	6	1	0	11	120	260
4	0	6	3	12	0	21	0	260
5	0	5	8	10	0	23	0	260
6	0	5	4	11	0	20	110	260
7	0	5	6	10	1	22	0	260
8	0	5	0	1	0	6	0	260
9	0	4	3	8	0	15	0	260
10	0	0	0	0	0	0	0	130
11	0	0	2	0	0	2	0	0
12	0	3	8	2	1	14	0	260
13	0	1	1	0	0	2	0	0
14	0	2	2	7	0	11	0	260
					Group Average	12.57143		
					St.Dev	7.713268		

PAPC2

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	5	2	11	0	18	120	130
2	0	6	2	11	0	19	120	260
3	1	6	8	1	0	16	120	260
4	1	6	3	12	0	22	120	260
5	1	6	8	11	0	26	0	260
6	1	6	8	12	0	27	110	260
7	1	5	6	11	1	24	0	260
8	1	6	2	11	0	20	120	260
9	0	4	6	11	0	21	0	260
10	0	0	0	11	0	11	0	130
11	1	5	2	7	0	15	120	260
12	0	3	8	12	1	24	120	260
13	0	3	1	12	0	16	0	80
14	0	4	2	12	0	18	120	260
					Group Average	19.78571		
					St.Dev	4.543441		

SNPC1

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	2	0	3	0	5	0	260
2	0	4	5	3	1	13	0	260
3	0	5	2	12	1	20	0	260
4	0	5	3	1	0	9	0	260
5	0	5	2	7	0	14	0	220
6	0	0	0	0	0	0	0	160
7	0	4	2	8	1	15	0	260
8	0	4	3	10	4	21	0	260
9	0	0	6	0	1	7	0	0
10	0	4	2	8	0	14	0	260
11	0	4	3	1	0	8	0	180
12	0	0	0	1	1	2	0	0
13	0	5	8	10	4	27	0	260
14	0	5	2	10	0	17	0	0
					Group Average	12.28571		
					St.Dev	7.620101		

SNPC2

ID	SignUp	Contact	EduInterest	Context	Marketing	Total	Connection	Settings
1	0	2	2	11	0	15	60	260
2	0	4	5	3	1	13	20	260
3	1	6	2	12	1	22	30	260
4	1	6	4	8	0	19	120	260
5	0	6	2	13	0	21	0	190
6	0	1	2	0	0	3	0	130
7	0	5	8	12	1	26	0	260
8	0	4	7	11	4	26	40	260
9	0	5	8	0	1	14	60	260
10	1	4	2	10	0	17	120	260
11	0	5	3	9	0	17	0	180
12	1	0	0	1	4	6	120	260
13	0	4	8	10	4	26	120	260
14	1	5	2	10	0	18	120	260
					Group Average	17.35714		
					St.Dev	6.979168		

Appendix 11 – Post-Experiment Interview Notes

The following is a summary of the notes taken during the post-experiment discussions.

Control Group

-Is it better to not answer questions, lie or is it OK to answer No?

With placement applications coming up in the next few years, I don't want any information about me to be out there so I think it is better to leave them blank.

-Did you?

No! I know I could have and I don't know why I answered them but now I wouldn't.

Some participants said lie is an appropriate method to protect themselves but would rather leave questions out.

-Do you need to protect your data if you're not answering questions?

I do it by habit, just see friends only and know that that is the one to pick.

You might end up submitting something in the future so should set it when you can.

-Why did you answer questions?

Because I could, I'm a "completionist". Felt good to be able to fill in all the fields.

Don't know, they were in front of me so I answered. Now we're talking about it I wouldn't have answered some of them.

-Connection settings

I didn't see the connection settings, I just went straight to the question bits.

I remember a link but didn't register it as important.

Personal Attitude

-What did you think of the lights?

They were useful in making me think about my privacy and what people will see and judge me on.

They did highlight things I wouldn't tell a stranger, like address.

It made me think about the information I put on social networks like twitter and Facebook and how that information could be used by others.

I get it, put I made up my own mind.

Made me think about what I want placement employees to see.

I didn't look at them but did leave things blank.

-Why did you leave things blank?

NTU doesn't need to know. I might tell my friends but not Facebook, there is no reason for it to know.

-Why did you answer things?

I answer the things that everyone is going to respond to so ("everyone does")

I left stuff out that I didn't want people to know the answer to.

-Do you need to protect even when not answering questions?

Yes, your profile is always changing, might be stuff there in the future.

Subjective Norms

-Do you need to worry about answering questions with a "No"?

Yes, I was embarrassed to answer no to some of them, people might think I am boring.

You are accountable for everything you put on there and the network could sell that information about you to someone else.

If the data is being view then leave it blank, people will use it judge you.

-Did you agree with the advice?

It reminded me to not answer everything but I don't remember following the advice.

I thought it might have been too strict.

Perceived Control

-Did you change your answers?

Yes, I wanted to get the score down – bit like a game.

Yes, it seemed to want me to.

-Better to not fill in questions or answer no (lie)?

Leave them out, if I have to answer them I will lie but if the choice is there.

-Do settings require application if no questions filled in?

Yes, I set everything to Friends Only.

You can't control what other people will post about you so it doesn't matter what you say.

PASN

-Did you prefer the lights or the advice?

The lights, they highlighted clearly what was bad without having to pay them too much attention.

The two together made me feel a bit anxious as there was too much to consider sometimes. Made the screen too busy.

-Need for privacy met?

Yes, I didn't answer what I thought was not relevant at the time.

A lot of this stuff is already on the web about me anyway, so I may as well fill it out again. If people want to find this information they can.

-Better to lie or withhold information?

Any data forms an opinion when others look at you.

-Still need to protect even when not answering questions?

Don't know, but I still put the highest for anything anyway.

Yes, things might change

PAPC

-Which did you prefer?

The lights; didn't need to review and go through all the extra screens.

-Did you change stuff then?

Yes, but not because I got anything wrong. Wanted low risk.

SNPC

-Did you change information?

Yes, but I'd already made my decision, didn't need to review but still made changes to alter the score.

It was overkill.

Just wanted to get the P-Score down.